

ULTIMATICS

Jurnal Teknik Informatika

RIKIP GINANJAR, ROSALINA, ALDO WIJAYA

Aplikasi Pemantauan Media Sosial untuk Analisa Merek

EKO WIYATNANTO, ARIEF INDRIARTO HARIS

Kinerja Arsitektur Interoperabilitas Menggunakan *Government Service Bus* (GSB) dan *Peer to Peer* (P2P)

RAHMAD AKBAR, BAMBANG PRAMONO, RIZAL ADI SAPUTRA

Implementasi Algoritma Simon Pada Aplikasi Kamus Perubahan Fi'il (Kata Kerja Bahasa Arab) Berbasis Android

MUHAMMAD NAUFAL RANDHIKA, JULIO CHRISTIAN YOUNG, ALETHEA SURYADIBRATA, HADIAN MANDALA

Implementasi Algoritma Complement dan Multinomial Naïve Bayes Classifier Pada Klasifikasi Kategori Berita Media Online

INSANUR HANIFUDDIN, DHOMAS HATTA FUDHOLI, SRI MULYATI

Sistem Konten Pembelajaran di Indonesia : *Systematic Literature Review*

DIMAS GALANG RAMADHAN, INDRI PRIHATINI, FEBRI LIANTONI

Analisis Clustering Pengelompokan Penjualan Paket Data Menggunakan Metode K-Means

TAUFIK KURNIALENSYA

Sistem Informasi Klinik Hewan Untuk Meningkatkan Kinerja Keuangan Menggunakan Metode Rasio Profitabilitas

DANIEL MARTOMANGGOLO WONOHADIDJOJO

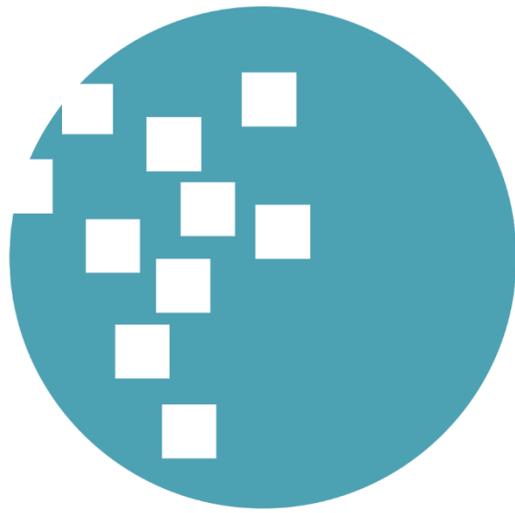
Perbandingan *Convolutional Neural Network* pada *Transfer Learning Method* untuk Mengklasifikasikan Sel Darah Putih

EUNIKE ENDARIAHNA S, ANDES SUCIANI, PHILIPUS SILAEN, SEPTIAN ADIWIBOWO

Pengenalan Aktivitas Manusia Melalui Analisis Data Gerakan *Smartphone*

M. FARIZ JANUARSYAH, ERMATITA

Prediksi Akurasi Kemenangan Pada Permainan Poker Menggunakan Algoritma C5.0 Dan WIPSO



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

EDITORIAL BOARD**Editor-in-Chief**

M.B.Nugraha, S.T., M.T.

Managing Editor

Ni Made Satvika Iswari, S.T., M.T.

Suryasari, S.Kom., M.T.

Eunike Endahriana Surbakti, S.Kom., M.T.I

Andre Rusli, S.Kom., M.Sc.

Nabila Rizky Oktadini, S.SI., M.T. (Unsri)

Rosa Reska Riskiana, S.T., M.T.I.

(Telkom University)

Hastie Audytra, S.Kom., M.T. (Unugiri)

Designer & Layouter

Intan Kusuma Wardani

Members

Rahmat Irsyada, S.Pd., M.Pd. (Unugiri)

Nirma Ceisa Santi (Unugiri)

Shinta Amalia Kusuma Wardhani (Unugiri)

Ariana Tulus Purnomo, Ph.D. (NTUST)

Alethea Suryadibrata, S.Kom., M.Eng. (UMN)

Dareen Halim, S.T., M.Sc. (UMN)

Fenina Adline Twince Tobing, M.Kom. (UMN)

Seng Hansun, S.Si., M.Cs. (UMN)

Dr. Moeljono Widjaja (UMN)

Wella, S.Kom., M.MSI., COBIT5 (UMN)

EDITORIAL ADDRESS

Universitas Multimedia Nusantara (UMN)

Jl. Scientia Boulevard

Gading Serpong

Tangerang, Banten - 15811

Indonesia

Phone. (021) 5422 0808

Fax. (021) 5422 0800

Email : ultimatics@umn.ac.id



Ultimatics : Jurnal Teknik Informatika is the Journal of the Informatics Study Program at Universitas Multimedia Nusantara which presents scientific research articles in the fields of Computer Science and Informatics, as well as the latest theoretical and practical issues, including Analysis and Design of Algorithm, Software Engineering, System and Network security, Ubiquitous and Mobile Computing, Artificial Intelligence and Machine Learning, Algorithm Theory, World Wide Web, Cryptography, as well as other topics in the field of Informatics. Ultimatics : Jurnal Teknik Informatika is published regularly twice a year (June and December) and is managed by the Informatics Study Program at Universitas Multimedia Nusantara.

Call for Papers



International Journal of New Media Technology (IJNMT) is a scholarly open access, peer-reviewed, and interdisciplinary journal focusing on theories, methods and implementations of new media technology. Topics include, but not limited to digital technology for creative industry, infrastructure technology, computing communication and networking, signal and image processing, intelligent system, control and embedded system, mobile and web based system, and robotics. IJNMT is published twice a year by Faculty of Engineering and Informatics of Universitas Multimedia Nusantara in cooperation with UMN Press.



Ultimatics : Jurnal Teknik Informatika is the Journal of the Informatics Study Program at Universitas Multimedia Nusantara which presents scientific research articles in the fields of Analysis and Design of Algorithm, Software Engineering, System and Network security, as well as the latest theoretical and practical issues, including Ubiquitous and Mobile Computing, Artificial Intelligence and Machine Learning, Algorithm Theory, World Wide Web, Cryptography, as well as other topics in the field of Informatics.



Ultima Computing : Jurnal Sistem Komputer is a Journal of Computer Engineering Study Program, Universitas Multimedia Nusantara which presents scientific research articles in the field of Computer Engineering and Electrical Engineering as well as current theoretical and practical issues, including Edge Computing, Internet-of-Things, Embedded Systems, Robotics, Control System, Network and Communication, System Integration, as well as other topics in the field of Computer Engineering and Electrical Engineering.



Ultima InfoSys : Jurnal Ilmu Sistem Informasi is a Journal of Information Systems Study Program at Universitas Multimedia Nusantara which presents scientific research articles in the field of Information Systems, as well as the latest theoretical and practical issues, including database systems, management information systems, system analysis and development, system project management information, programming, mobile information system, and other topics related to Information Systems.

KATA PENGANTAR

Salam ULTIMA!

Ultimatics : Jurnal Teknik Informatika merupakan Jurnal Program Studi Informatika Universitas Multimedia Nusantara yang menyajikan artikel-artikel penelitian ilmiah dalam bidang Informatika serta isu-isu teoritis dan praktis yang terkini, mencakup Analysis and Design of Algorithm, Software Engineering, System and Network security, Ubiquitous and Mobile Computing, Artificial Intelligence and Machine Learning, Algorithm Theory, World Wide Web, Cryptography, serta topik lainnya di bidang Informatika.

Pada edisi Juni 2021 ini, ULTIMA Computing memasuki Edisi ke-1 dari Volume 13. Dalam jurnal ini terdapat enam buah karya tulis ilmiah yang berasal dari para peneliti, akademisi dan praktisi di bidang Informatika dan Ilmu Komputer. Beberapa topik yang diangkat dalam jurnal ini, antara lain Aplikasi Pemantauan Media Sosial untuk Analisa Merek, Kinerja Arsitektur Interoperabilitas Menggunakan *Government Service Bus* (GSB) dan *Peer to Peer* (P2P), Implementasi Algoritma Simon Pada Aplikasi Kamus Perubahan Fi'il (Kata Kerja Bahasa Arab) Berbasis Android, Implementasi Algoritma Complement dan Multinomial Naïve Bayes Classifier Pada Klasifikasi Kategori Berita Media Online, Sistem Konten Pembelajaran di Indonesia : *Systematic Literature Review*, Analisis Clustering Pengelompokan Penjualan Paket Data Menggunakan Metode K-Means, Sistem Informasi Klinik Hewan Untuk Meningkatkan Kinerja Keuangan Menggunakan Metode Rasio Profitabilitas, Perbandingan *Convolutional Neural Network* pada *Transfer Learning Method* untuk Mengklasifikasikan Sel Darah Putih, Pengenalan Aktivitas Manusia Melalui Analisis Data Gerakan *Smartphone*, dan Prediksi Akurasi Kemenangan Pada Permainan Poker Menggunakan Algoritma C5.0 Dan WIPSO.

Pada kesempatan kali ini juga kami ingin mengundang partisipasi para pembaca yang budiman, para peneliti, akademisi, maupun praktisi, di bidang Teknik dan Informatika, untuk mengirimkan karya ilmiah yang berkualitas pada: International Journal of New Media Technology (IJNMT), Ultimatics : Jurnal Teknik Informatika, Ultima Infosys : Jurnal Sistem Informasi dan Ultima Computing : Jurnal Sistem Komputer. Informasi mengenai pedoman dan template penulisan, serta informasi terkait lainnya dapat diperoleh melalui alamat surel ultimatics@umn.ac.id dan halaman web dari Jurnal kami di sini. Perlu diketahui, mulai saat ini kami hanya akan menerima dan memproses artikel yang ditulis dalam Bahasa Inggris sebagai tindak lanjut dari upaya migrasi kami ke jurnal berbahasa Inggris.

Akhir kata, kami mengucapkan terima kasih kepada seluruh kontributor dalam Ultimatics Edisi Juni 2021 ini. Kami berharap artikel-artikel ilmiah hasil penelitian dalam jurnal ini dapat bermanfaat dan memberikan sumbangsih terhadap perkembangan penelitian dan keilmuan di Indonesia.

Juni 2021,

M.B.Nugraha, S.T., M.T.
Ketua Dewan Redaksi

DAFTAR ISI

Aplikasi Pemantauan Media Sosial untuk Analisa Merek Rikip Ginanjar, Rosalina, Aldo Wijaya	1-6
Kinerja Arsitektur Interoperabilitas Menggunakan <i>Government Service Bus</i> (GSB) dan <i>Peer to Peer</i> (P2P) Eko Wiyatnanto, Arief Indriarto Haris	7-11
Implementasi Algoritma Simon Pada Aplikasi Kamus Perubahan Fi'il (Kata Kerja Bahasa Arab) Berbasis Android Rahmad Akbar, Bambang Pramono, Rizal Adi Saputra	12-18
Implementasi Algoritma Complement dan Multinomial Naïve Bayes Classifier Pada Klasifikasi Kategori Berita Media Online Muhammad Naufal Randhika, Julio Christian Young, Alethea Suryadibrata, Hadian Mandala	19-25
Sistem Konten Pembelajaran di Indonesia : <i>Systematic Literature Review</i> Insanur Hanifuddin, Dhomas Hatta Fudholi, Sri Mulyati	26-32
Analisis <i>Clustering</i> Pengelompokan Penjualan Paket Data Menggunakan Metode K-Means Dimas Galang Ramadhan, Indri Prihatini, Febri Liantoni	33-38
Sistem Informasi Klinik Hewan Untuk Meningkatkan Kinerja Keuangan Menggunakan Metode Rasio Profitabilitas Taufik Kurnialensya	39-50
Perbandingan <i>Convolutional Neural Network</i> pada <i>Transfer Learning Method</i> untuk Mengklasifikasikan Sel Darah Putih Daniel Martomanggolo Wonohadidjojo	51-57
Pengenalan Aktivitas Manusia Melalui Analisis Data Gerakan <i>Smartphone</i> Eunike Endariahna S, Andes Suciani, Philipus Silaen, Septian Adiwibowo	58-65
Prediksi Akurasi Kemenangan Pada Permainan Poker Menggunakan Algoritma C5.0 Dan WIPSO M. Fariz Januarsyah, Dr. Ermatita, M.Kom.	66-71

Aplikasi Pemantauan Media Sosial untuk Analisa Merek

Rikip Ginanjar¹, Rosalina², Aldo Wijaya³

^{1,2,3} Faculty of Computing, President University, Bekasi, Indonesia

¹rosalina@president.ac.id

Diterima 02 September 2020

Disetujui 26 November 2020

Abstract— In recent years, micro-blogs on the Internet have become a popular way of expressing feelings, thoughts, and even communicating opinions about products and services that are common among its users. Collecting user opinions can be an expensive and time-consuming task using conventional methods such as surveys. The sentiment analysis of the customer opinions makes it easier for businesses to understand their competitive value in a changing market and to understand their customer views about their products and services. In this research, Lexicon-Based approach especially AFINN lexicon is implemented to classify user twitter sentiment, throughout which, twitter Micro-blogs data has been collected, pre-processed analyzed, and classified. The results of this research is an android application that could classify users' perspective via tweets into positive and negative, which is represented in a pie chart for Monthly report.

Index Terms— Sentiment Analysis, Brand Analysis, Twitter, Android Application

I. PENDAHULUAN

Media sosial memberikan data yang bermanfaat bagi sumber pembelajaran, dimana secara statistik [1] diverifikasi bahwa 71 *web* digunakan oleh para konsumen dengan menggunakan media sosial. Terkait dengan hal tersebut, penelitian [2] menunjukkan bahwa 50% konsumen memilih untuk melihat komentar para konsumen lainnya tentang produk sebelum memutuskan untuk memberikan produk atau layanan tersebut. Pendapat para konsumen merupakan umpan balik yang paling efektif tentang produk dan layanan bagi para pelaku bisnis yang memungkinkan untuk mengubahnya menjadi peluang bisnis baru [3]. Jejaring sosial seperti *Twitter* menyediakan media promosi, *branding* serta meningkatkan promosi merek dagang [4].

Sentimen adalah sikap, pemikiran, atau penilaian yang didorong oleh perasaan [5]. Analisis sentimen didefinisikan sebagai proses penggalian dan pemahaman sentimen yang diungkapkan dalam dokumen teks. Analisis sentimen, juga dikenal sebagai penambangan opini [6,7]. Dari sudut pandang pengguna, mereka dapat memposting kontennya sendiri melalui berbagai media sosial, seperti forum, mikro-blog, atau situs jejaring sosial *online* salah satunya *Twitter* [8]. Dari perspektif peneliti, banyak

situs media sosial merilis antar muka pemrograman aplikasi (API), yang membantu para peneliti dalam pengumpulan dan analisis data, misalnya, *Twitter* saat ini memiliki beberapa API yang tersedia, yaitu *SEARCH API*. Dengan *SEARCH API*, *developer* dapat mengumpulkan data status dan informasi pengguna untuk dijadikan bahan penelitian.

Situs *web micro-blogging* telah berkembang menjadi sumber berbagai jenis informasi. Hal ini disebabkan oleh sifat mikro-blog yang merupakan tempat orang atau para penggunanya untuk memposting pesan mengenai pendapat mereka mengenai berbagai topik, mendiskusikan masalah terkini, mengeluh, serta mengungkapkan sentimen positif untuk merek yang mereka gunakan dalam kehidupan sehari-hari. Perusahaan yang memproduksi merek semacam itu sudah mulai menganalisis blog mikro ini untuk mengetahui sentimen masyarakat mengenai merek dagang mereka. Sering kali, perusahaan-perusahaan ini mempelajari reaksi pengguna dan membalas pengguna di blog mikro. Terkait dengan ini, salah satu tantangannya adalah membangun teknologi untuk mendeteksi dan meringkas sentimen secara keseluruhan. Dalam penelitian ini, dilakukan analisa *tweet* dari *Twitter* mengenai opini masyarakat mengenai merek dagang tertentu.

Tweet adalah sumber informasi yang dapat diandalkan terutama karena penggunaannya men-*tweet* tentang apa saja dan semua yang mereka lakukan termasuk membeli merek barang atau jasa tertentu dan kemudian mengulasnya. Selain itu, semua *tweet* berisi *hashtag* yang memudahkan dalam mengidentifikasi *tweet* dibandingkan dengan menggunakan metode konvensional yang menghabiskan waktu yang banyak terutama dalam menganalisa dan mengumpulkan opini pengguna. Sejumlah penelitian mengenai sentiment analysis terhadap merek dagang dilakukan dengan menggunakan *Twitter* [9-14], penelitian-penelitian tersebut berhasil mengidentifikasi opini pengguna yang berpengaruh dalam branding serta meningkatkan promosi merek dagang.

AFINN telah dikembangkan oleh Finn Arup Nilesen sebagai metode analisis sentimen khusus untuk *microblogging*, AFINN adalah metode Berbasis Leksikon yang berisi daftar kata-kata yang dinilai untuk valensi dengan skor *integer*, metode leksikon

AFINN adalah metode Leksikon paling sederhana dan paling populer yang dapat digunakan secara luas untuk analisis sentimen. Aplikasi tersebut mengumpulkan data menggunakan API publik Twitter yang memungkinkan pengembang mengekstrak *tweet* dari Twitter secara pro-gramatikal. Data yang dikumpulkan, karena sifat *tweeting* yang acak dan kasual, perlu disaring untuk menghapus informasi yang tidak perlu seperti menghapus nama pengguna, *hashtag*, *URL*, dan lain sebagainya. Karena tahap *preprocessing* telah dilakukan sampai batas tertentu, mungkin untuk menjamin bahwa analisis *tweet* yang difilter ini akan memberikan hasil yang dapat diandalkan.

Adapun tujuan utama dari penelitian ini adalah membuat aplikasi *android* yang dapat menganalisis *tweet*, membuat daftar *brand* sesuai kategorinya dan divisualisasikan dalam bentuk *bar chart* sesuai dengan nilai rating tertinggi, mengimplementasikan AFINN Lexicon untuk klasifikasi teks ke dalam kategori positif, dan negatif, dan membuat grafik persentase sentimen setiap bulan dari data *tweet* terbaru. Batasan penelitian ini adalah sebagai berikut:

- Aplikasi dibuat untuk menentukan opini berdasarkan *tweet* hanya dalam bahasa Inggris.
- Aplikasi yang dikembangkan hanya menganalisis kurang dari 5000 *tweet* sesuai ketersediaan data.
- Hanya menyajikan beberapa kategori *brand* atau merek yang dapat dianalisis.

II. LANDASAN TEORI

A. Sentiment Analysis

Suatu instansi secara konsisten harus mengetahui penilaian pembeli mengenai produk dan layanan mereka. Begitu juga dengan pembeli atau konsumen perlu mengetahui penilaian pembeli lain sebelum mereka memutuskan untuk membeli barang atau menggunakan suatu layanan. Analisis sentimen merupakan salah satu pendekatan yang digunakan untuk menganalisis opini positif, negatif dan netral orang tentang merek atau jasa tertentu [15]. Analisis Sentimen tidak hanya bergantung pada polaritas kata, tetapi juga pada pohon tata bahasa dari kalimat yang dianalisis. Sistem mengidentifikasi ekspresi idiomatik, memberikan interpretasi pada negasi, memodifikasi polaritas kata yang berdasarkan pada adverbia, kata sifat, konjungsi atau kata kerja yang terhubung, terutama dengan mempertimbangkan peningkatan logika fungsional tertentu.

Pesan Twitter kerap digunakan untuk menentukan sentimen pembeli terhadap suatu merek. Penelitian terkini tentang analisis sentimen di Twitter adalah sumber jangkauan data sosial yang merupakan titik awal untuk penambangan web sosial karena keterbukaannya yang melekat dengan konsumsi publik, API yang terdokumentasi dengan baik, serta daya tarik yang luas bagi pengguna dari setiap lapisan masyarakat [15].

Dalam beberapa tahun terakhir banyak teknik analisis sentimen yang telah dikembangkan. Pertama, klasifikasi dilakukan untuk mengklasifikasikan teks menjadi kelas sentimen positif atau negatif. Kemudian dilakukan pendekatan esensial berbasis leksikon guna menyelidiki *tweet* yang terkandung di dalam teks. Teks dipindai dan diperiksa untuk mengetahui apakah terdapat kata sentimental yang terkandung didalamnya. Setelah itu, kata-kata tersebut kemudian diberi nilai sentimen yang ditentukan berdasarkan nilai atau skor.

B. AFINN Lexicon

Metode berbasis leksikon disebut juga sebagai metode leksikon dan berfokus pada korpus atau kamus kata-kata dengan polaritas yang telah ditentukan. Leksikon AFINN dibuat pada tahun 2009 untuk *tweet* yang diunduh untuk analisis sentimen *online*. Versi sebelumnya AFINN-96 berisi 1468 kata-kata yang sangat berbeda. Versi terbaru yang dikenal sebagai AFINN-111 berisi 2477 kata. Skor di AFINN mulai dari -5 hingga +5. -5 berarti sangat negatif dan +5 berarti sangat positif. Kata-kata tersebut dinilai secara manual. Glosarium dalam leksikon AFINN dimulai dengan kumpulan kata-kata negatif [16]. Sebagian besar kata positif diberi skor dengan +2 dan sebagian besar kata negatif diberi skor -2, kata-kata negatif yang kuat dengan nilai -4 atau -5.

C. Penelitian terkait

SentiTweet [17] merupakan aplikasi yang memberikan wawasan atau informasi mengenai kata-kata yang telah digunakan di *tweet* terbaru dan juga untuk memprediksi suasana hati seseorang berdasarkan 20 *tweet* teratas di *timeline*-nya. Ketika pengguna memasukkan kata kunci, hasilnya hanya menampilkan Diagram Lingkaran dengan 10 *tweeter* teratas, dan 10 lokasi *tweet* teratas tanpa memberikan hasil seperti detail siapa yang telah *tweet*, persentase hanya mengandalkan Diagram Lingkaran.

Oleh karena itu penelitian ini bertujuan untuk membuat informasi yang lebih detail seperti siapa yang telah *tweet*, *tweet* tersebut kemudian hasilnya dibagi menjadi 2 kategori yaitu Positive, dan Negative Sentiment serta untuk menambahkan *Pie Chart* (untuk menunjukkan persentase hasil Sentiment tersebut di tiap bulannya), antarmuka pengguna *SentiTweet* ditunjukkan pada Gambar 1.



Gambar 1. Tampilan SentiTweet [17]

Tabel 1. Perbandingan antara penelitian dan penelitian terkait

Perbandingan	Sentisis (Penelitian ini)	TweetMent	SentiTweet
Daftar Merek	V	X	X
Bar chart merek berdasarkan rating	V	X	X
Deskripsi dari merek	V	X	X
Hasil detail bulanan untuk setiap bulan	V	X	X
Grafik persentase Sentimen	V	V	V
Daftar tweet berdasarkan klasifikasi sentiment	V	V	X

TweetMent [18] adalah aplikasi seluler yang mengandalkan API *Twitter* dan pengklasifikasi *Naive Bayes* untuk mengklasifikasikan *tweet* yang diposting (maksimal 7 hari) dan memberikan sentimen umum pengguna pada topik pencarian tertentu. Aplikasi ini dapat mencari *tweet* (*tweet* apapun), hasilnya berupa daftar *tweet* dengan analisis sentiment; *tweet* negatif dengan warna merah, *tweet* positif dengan warna hijau, dan *tweet* netral dengan warna abu-abu, masing-masing *tweet* menampilkan *username*, foto pengguna, dan *tweet* mereka serta hasil pencarian lainnya yang divisualisasikan pada grafik berdasarkan total sensitivitas, tampilan antar muka *TweetMent* terlihat pada Gambar 2.



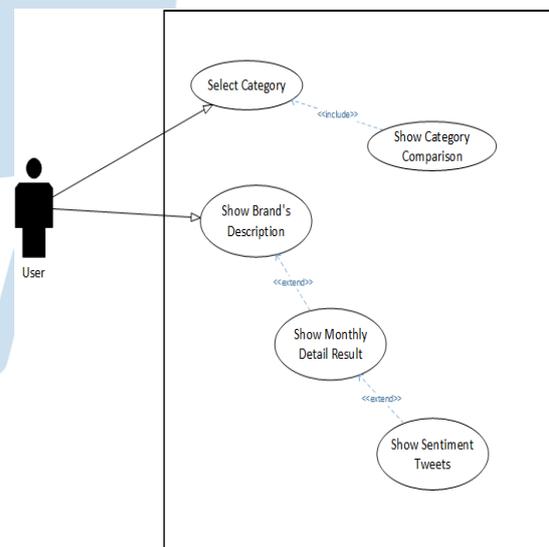
Gambar 2. Tampilan Antarmuka TweetMent [18]

Perbandingan antara penelitian dan penelitian terkait ditunjukkan pada Tabel 1.

III. METODE PENELITIAN

Metode penelitian Aplikasi Pemantauan Media Sosial untuk Analisa Merek dilakukan dalam beberapa tahapan seperti data teks atau pengumpulan data, data *preprocessing*, *tokenization*, *AFINN Scoring*, Klasifikasi *Tweet*.

Aplikasi pemantauan media sosial untuk analisa merek merupakan aplikasi berbasis Android yang digunakan untuk menganalisis sentimen *tweet*. Fitur utama dari aplikasi ini adalah melakukan analisis sentimen dalam *tweet* dan mengklasifikasikannya menjadi sentimen positif yang ditunjukkan dengan warna hijau, dan sentimen negatif yang ditunjukkan dengan warna merah. Aplikasi ini mengimplementasikan pendekatan *Lexicon-Based* khususnya *AFINN lexicon* untuk mengklasifikasikan sentimen *twitter* pengguna. Data *tweet* divisualisasikan dengan *Pie Chart* sebagai ringkasan dari total sentimen *tweet* di setiap bulannya. Aplikasi android ini dapat digunakan oleh pengguna, organisasi, perusahaan untuk mereview citra merek, adapun diagram *use case* dari aplikasi ini ditunjukkan pada Gambar 3.



Gambar 3. Diagram Use Case

A. Tweet Search API Dataset

Berdasarkan pencarian merek, aplikasi ini bisa mendapatkan kurang dari 5000 *tweet* (juga tergantung jumlah data yang tersedia) menggunakan *Twitter Search API*. Sedangkan selama proses pencarian aplikasi ini memfilter data mentah dan menghapus semua *tweet* yang dianggap *noise*. Sebuah *tweet* dianggap berisik jika berisi "RT", "@", "URL", dan lain sebagainya.

B. Twitter API

Pada aplikasi ini untuk mengambil data dengan *Twitter Search API* dilakukan dengan menggunakan metode *request* and *response*, responnya berupa format JSON, dengan syarat sekali *request* tidak boleh lebih dari 180 *tweet*, jika melebihi maka akan mempengaruhi rate limit. Terkait dengan hal tersebut, diperlukan permintaan perulangan untuk mendapatkan lebih dari 180 *tweet*, adapun konten JSON didapatkan dari status pengguna *twitter* seperti *create_at*, teks, url, metadata, dan lain sebagainya.

C. Tweet Preprocessing

Tahapan *pre-processing* memodifikasi *tweet* mentah sebelum diteruskan ke fitur ekstraksi. Selain itu juga merupakan tahapan yang diperlukan untuk meningkatkan kinerja sistem secara keseluruhan. Dalam tahap ini metode sederhana dirangkai bersama untuk memodifikasi *tweet* mentah menggunakan ekspresi reguler. Membatasi agar setiap *pre-processor* hanya melakukan satu tugas sederhana sehingga mudah dalam pengelolaannya seperti yang terlihat pada Tabel 2.

Tabel 2. Daftar *preprocessor* yang digunakan dalam *Tweet Preprocessing*

Filter	Deskripsi
Lower Case	Mengubah seluruh huruf besar menjadi huruf kecil
No User	Mengganti seluruh <i>username</i> yang disebutkan dengan <i>empty string</i>
No RT Tag	Mengganti seluruh <i>tags</i> RT dengan <i>empty string</i>
No URL	Mengganti seluruh <i>tags</i> URL dengan <i>empty string</i>
No Hashtag	Mengganti semua tanda pagar bersama dengan tag berikut dengan string kosong
Limit Chars	Menghapus semua karakter non abjad atau spasi

D. AFINN Lexicon

Pendekatan berbasis leksikon disebut juga pendekatan kamus dan mengandalkan leksikon atau kamus kata dengan polaritas yang telah dihitung sebelumnya. Kadang-kadang metode ini dianggap sebagai bagian dari pendekatan *Machine Learning Unsupervised*, namun juga menggambarkannya metode independen, karena kualitas klasifikasi dalam pendekatan berbasis leksikon hanya bergantung pada kualitas leksikon. Pendekatan berbasis leksikon sebagai berikut:

1. Membangun leksikon kata dan kata-kata tersebut diberi nilai

2. Menyiapkan *bag of words* yang telah dibuat sebelumnya (AFINN) yang akan digunakan dalam memberikan nilai sebuah kata
3. Tahap *preprocessing*: penurunan kata dalam huruf besar / kecil, penghilangan kata-kata penghenti, stemming, negasi.
4. Skor perhitungan sentimen: setiap kata dari sekumpulan kata akan dibandingkan dengan leksikon AFINN. Jika kata ditemukan dalam leksikon, skor sentimen kata tersebut kemudian ditambahkan dan skor sentimen dijumlahkan.

Contohnya:

“A masterful[+0.92] film from a master[+1] filmmaker ,unique[+1] in its deceptive grimness , compelling[+1] in its fatalist[-0.84] worldview.”

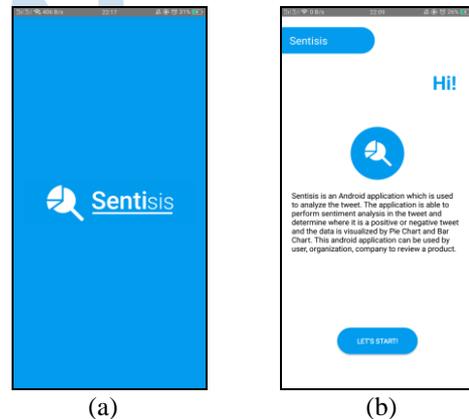
Akan dihitung sebagai berikut:

Total Sentiment Score = +0.92 +1 +1 +1 -0.84 = 3.08
Yang berarti, teks tersebut mengandung sentimen positif.

IV. HASIL PENELITIAN

A. User Interface Development

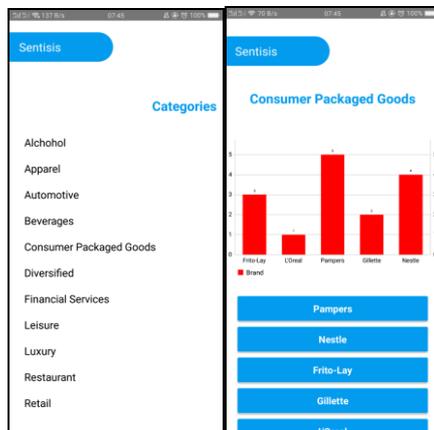
Saat pertama kali digunakan gambar 4 akan ditampilkan, gambar tersebut akan muncul saat pengguna mengklik ikon aplikasi. Pada layar akan terlihat halaman dengan logo aplikasi, dan setelah 3 detik, layar ini akan digantikan oleh halaman beranda. Sedangkan gambar 4(b) menunjukkan *home* atau halaman depan aplikasi, pada layar ini terdapat *TitleBar* dengan nama aplikasi, teks untuk menyapa pengguna, logo aplikasi dan deskripsi aplikasi, dan " Lets Start" untuk menuju ke layar kategori.



Gambar 4. (a) *Splash Screen*, (b) Halaman Beranda

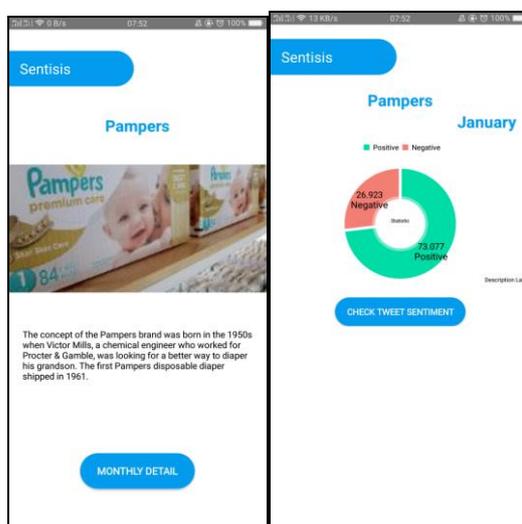
Gambar 5(a) menampilkan halaman kategori, didalamnya terdapat beberapa kategori merek yang telah diinput secara manual, dan adapun sumber input diambil dari Majalah Forbes. Pada halaman ini, pengguna dapat memilih merek yang akan dianalisis dan pengguna juga dapat melihat perbandingan merek di grafik di layar berikutnya. Sedangkan gambar 5(b) menampilkan halaman perbandingan dengan grafik

batang. Perbandingan merek berdasarkan jumlah peringkat dimana 5 adalah peringkat tertinggi dan 1 adalah peringkat terendah dan juga pada halaman tersebut terdapat tombol Nama Merek. Pengguna dapat mengklik tombol Nama Merek yang ingin dianalisis.



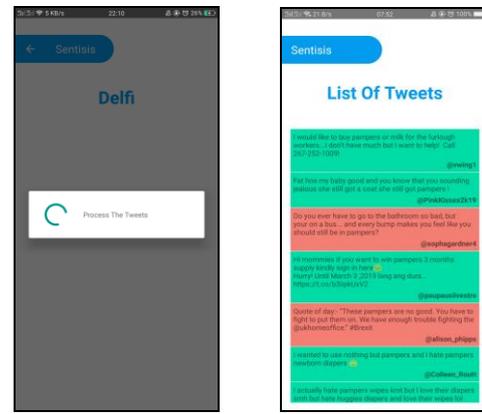
Gambar 5. (a) Halaman kategori, (b) Halaman grafik

Gambar 6(a) menunjukkan halaman hasil ketika pengguna memilih merek yang dipilih. Pada halaman ini ditampilkan nama merek, gambar/logo merek mereka, juga terdapat penjelasan singkat dari merek untuk memberi informasi ke pengguna mengenai jenis merek yang dipilih, selain itu, juga terdapat tombol untuk menampilkan detail bulanan di layar berikutnya. Sedangkan gambar 6(b) menunjukkan halaman detail ketika pengguna menekan tombol "Monthly Detail" di halaman deskripsi, di halaman ini ditampilkan diagram lingkaran, dimana diagram lingkaran tersebut berisi persentase total sentimen positif dan negatif, dan terdapat tombol "Periksa Tweet" untuk memeriksa tweet positif dan negatif dari pengguna di twitter berdasarkan bulan yang dipilih.



Gambar 6. (a) Halaman hasil, (b) Halaman detail

Gambar 7(a) dan 7(b) menunjukkan tampilan aplikasi ketika pengguna mengklik tombol "Check Tweet", tombol tersebut akan membawa pengguna menuju ke layar tweet. Daftar tweet berisi sentimen positif dan negatif, teks dengan warna merah berarti sentimen negatif dan teks dengan warna hijau berarti sentimen positif.



Gambar 7. (a) Loading Screen, (b) Tweet Screen

V. KESIMPULAN

Penelitian ini mengimplementasikan metode *Lexicon-Based* untuk mengklasifikasikan tweet. Aplikasi yang dikembangkan dapat menentukan sentimen positif dan negatif dari setiap tweet. Berdasarkan hasil tersebut, kemudian hasilnya direpresentasikan dalam diagram lingkaran sebagai rangkuman dari total tweet sentimen. Aplikasi yang dikembangkan sesuai dengan tujuan dari penelitian ini, namun aplikasi ini kedepannya tetap perlu dikembangkan, dengan pertimbangan bahwa semua jenis aplikasi harus dipelihara dan selalu diperbarui dengan teknologi terkini agar fungsi dan kebutuhan pengguna selalu terjaga. Berikut adalah rekomendasi untuk pengembangan dimasa datang :

- Agar aplikasi dapat bekerja pada platform yang berbeda (lintas platform) sehingga aplikasi dapat berjalan tidak hanya pada sistem operasi Android, tetapi aplikasi dapat digunakan pada beberapa sistem operasi yang berbeda seperti Microsoft Windows, Linux, Mac OS, dan lain sebagainya.
- Aplikasi tidak hanya menganalisis dataset sebuah tweet dari Twitter tetapi juga dapat mengambil dataset sentimen dari media sosial lain seperti Facebook, Instagram, dan situs berita online seperti Kompas, Detik, OkeZone, LineToday dan lain sebagainya.
- Dalam aplikasi ini tweet yang diambil kurang dari 5000 tweet. Dengan pengambilan data tweet lebih dari 5000 tweet, akan mempengaruhi rentang bulan yang didapat menjadi lebih luas sehingga data dan detail yang dikumpulkan akan lebih bertambah. Tambahkan lebih banyak

kategori merek yang akan dianalisis agar datanya lebih banyak dan beragam serta penambahan fitur pencarian dan saran pencarian agar lebih mudah dan cepat dalam pencarian kategori untuk pengembangan lebih lanjut.

DAFTAR PUSTAKA

- [1] Statista.com, “number of social media users worldwide from 2010 to 2021 in Billions, Link: <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>
- [2] DoubleClick_Inc .Understanding Buyer Search Activity as it Builds to Online Purchase. Retrieved from: www.innovationmarketing.at/news/newsmodul/upload/429181378_searchpurchase_0502.pdf. 2005
- [3] Hu, G., Bhargava, P., Fuhrmann, S., et al: ‘Analyzing users’ sentiment towards popular consumer industries and brands on twitter’. IEEE Int. Conf. on Data Mining Workshops, ICDMW, New Orleans, LA, USA., pp. 381–388. 2017
- [4] Sharma and Alavi, “Generating trust using Facebook-A study of 5 online apparel brands”, Information Technology and Quantitative Management (ITQM2017), 2017.
- [5] Kim S-M, Hovy E, “Determining the sentiment of opinions”, Proceedings of the 20th international conference on Computational Linguistics, page 1367. Association for Computational Linguistics, Stroudsburg, PA, USA, 2004.
- [6] Liu B, “Sentiment analysis and subjectivity:, In: Handbook of Natural Language Processing, Second Edition, 2010.
- [7] Taylor and Francis Group, Boca. Liu B, Hu M, Cheng J, “Opinion observer: Analyzing and comparing opinions on the web”, In: Proceedings of the 14th International Conference on World Wide Web, WWW ’05, 342–351. ACM, New York, NY, USA, 2005.
- [8] Pak A, Paroubek P (2010) Twitter as a corpus for sentiment analysis and opinion mining In: Proceedings of the Seventh conference on International Language Resources and Evaluation. European Languages Resources Association, Valletta, Malta, 2010.
- [9] Haripriya V., Patil P.G, “Detection of Sarcasm from Consumer Sentiments on Social Media About Luxury Brands”, In: Santosh K., Hegadi R. (eds) Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2018. Communications in Computer and Information Science, vol 1037. Springer, Singapore, 2019
- [10] Bigonha, C., Cardoso, T. N. C., Moro, M. M., et al: ‘Sentiment-based influence detection on twitter’, J. Braz. Comput. Soc., 18, (3), pp. 169–183, 2012
- [11] Hamid Bagheri, Md Johirul Islam, “Sentiment analysis of twitter data”, Annual International Conference “Dialogue”, pp. 14-28, 2017
- [12] Aliza Sarlan, Chayanit Nadam and Shuib Basri, “Twitter Sentiment Analysis”, 2014 International Conference on data Technology and Multimedia (ICIMU), Putrajaya, Malaya Gregorian calendar month eighteen – twenty, 2014.
- [13] Hasan, A., Moin, S., Karim, A., et al: ‘Machine learning-based sentiment analysis for twitter accounts’, Math. Comput. Appl., 2017, 5, pp. 3008–3015, 2017
- [14] Bharti, S.K., Vachha, B., Pradhan, R.K., Babu, K.S., Jena, S.K.: Sarcastic sentiment detection in tweets streamed in real time: a bigdata approach. Digit. Commun. Netw. <http://dx.doi.org/10.1016/j.dcan.2016.06.002>
- [15] Umman Tugba Gursoy, Diren Bulut, Cimel Yigit. “Social Media Mining and Sentiment Analysis for Brand Management”, Globat Journal of Emerging Trends in E-Business, Marketing and Consumer Psychology (GJETeMCP), Vol. 3, Issue 1. 2017
- [16] Finn Arup Nielsen, "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs", ESWC2011 Workshop on Making Sense of Micro posts: March 2011, pp. 93-98., 2011.
- [17] SentiTweet, Link: <http://www.sentitweet.com>, last accessed on: June 16, 2020
- [18] TweetMent, Appstore Link: <https://play.google.com/store/apps/details?id=org.junejadipan.shu.tweetment&hl=en>, last accessed: June 16, 2020



Kinerja Arsitektur Interoperabilitas Menggunakan *Government Service Bus* (GSB) dan *Peer to Peer* (P2P)

Eko Wiyatnanto¹, Arief Indriarto Haris²

^{1,2}Pusat Teknologi Informasi dan Komunikasi Penerbangan dan Antariksa, Lembaga Penerbangan dan Antariksa Nasional (LAPAN), Jakarta, Indonesia

¹eko.wiyatnanto@lapan.go.id, ²arief.indriarto@lapan.go.id

Diterima 05 November 2020

Disetujui 24 Mei 2021

Abstract — The role of interoperability architecture is one of the solutions to data redundancy problems and data differences that cause the level of data accuracy to be low. However, the performance of the interoperability architecture needs to be evaluated as an effort to improve the quality of the application itself. This study aims to evaluate the interoperability architecture between architectures using the Government Service Bus (GSB) and Peer to Peer (P2P), through several tests, namely load testing and stress testing. Load testing and stress testing aim to measure the speed and resilience of an application by sending requests and measuring the response of the application. The difference is that in load testing, testing is carried out with certain load conditions, while in stress testing, testing is carried out under extreme conditions. Testing load testing, determining the load condition with a multiple of the number of users 100 to 500, with the JMeter tools. In stress testing, testing is carried out with a total of 1000 users, with the Loader.io tools. The results of the load testing show that the average access time for GSB is smaller than P2P, with an average of 2ms to 309ms and a request error rate of 0%. The results of stress testing show that the GSB architecture is faster than the P2P architecture with an average difference in access time of 2957ms, a difference in throughput of 1.5 request/second, but GSB has a higher request error rate than P2P with a difference of 16.05%. In general conditions with certain loads, the GSB architecture has superior performance, this can be seen from the access time and request errors. Meanwhile, in extreme conditions the GSB architecture experiences a decline in performance, this can be seen from a larger request error rate when handling very large requests.

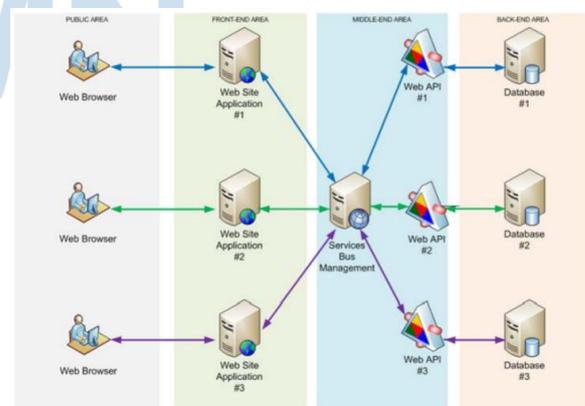
Index Terms—Government Service Bus; Interoperability Architecture; MANTRA

I. PENDAHULUAN

Interoperabilitas *data* aplikasi di lingkungan Lembaga Penerbangan dan Antariksa Nasional (LAPAN) sangat dibutuhkan untuk mendorong efisiensi serta penghematan biaya pengembangan aplikasi. LAPAN memiliki aplikasi dengan berbagai bahasa pemrograman dan framework aplikasi yang berbeda-beda dan hampir setiap aplikasi

dikembangkan menggunakan DBMS (*Database Management System*). Pembangunan aplikasi terus dilakukan tanpa interkoneksi data mengakibatkan terjadi redudansi *data* dan perbedaan *data* sehingga keakuratan *data* menjadi rendah. Hal ini mengakibatkan ketersediaan *data* tidak optimal. Oleh karena itu perlu dibangun sistem yang mampu membuat interkoneksi sebagai media pertukaran *data* dan informasi antar sistem.

Secara sederhana interoperabilitas dapat dikatakan sebagai suatu sistem yang memiliki kemampuan berkomunikasi dan bekerja sama dengan sistem lain tanpa adanya batasan terhadap akses informasi [1]. *Government Service Bus* (GSB) merupakan suatu sistem yang mengelola integrasi informasi dan membentuk interkoneksi sebagai jembatan pertukaran data antar sistem instansi pemerintah menggunakan *Application Programming Interface* (API) [2].

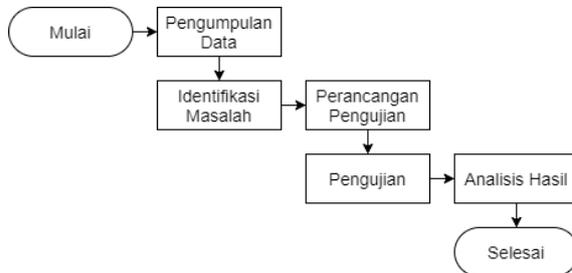


Gambar 1. Topologi Layanan API/Webservice melalui GSB [2]

Pemanfaatan akses layanan berbagi pakai menggunakan API *Webservice Point to Point* (P2P) antar *server* membentuk koneksi komunikasi sebagai media lalu lintas *data* [2].

II. METODOLOGI PENELITIAN

Pada penelitian ini, terdapat serangkaian tahapan yang dilakukan (Gambar 5). Pada tahap pengumpulan data, penulis melakukan observasi langsung terkait kebutuhan sistem informasi di lingkungan LAPAN, studi literatur, dan wawancara terhadap penyedia layanan dan beberapa *stakeholder*. Setelah dilakukan pengumpulan data, penulis melakukan identifikasi masalah terkait kinerja arsitektur interoperabilitas menggunakan GSB dan P2P.



Gambar 6. Tahapan Penelitian

Pada tahap perancangan pengujian, penulis melakukan perancangan terkait metode pengujian, skenario pengujian, hingga *tools* dan *parameter* yang akan digunakan, serta penyiapan instrumen pendukung lainnya seperti penyiapan *endpoint* sebagai target data. Pada tahap pengujian untuk mengukur kinerja GSB dan P2P, penulis menggunakan metode *load testing* dan *stress testing*. *Load testing* dan *stress testing*, keduanya memiliki kesamaan tujuan yaitu untuk mengukur kecepatan dan ketahanan suatu aplikasi dengan cara mengirimkan *request* dan mengukur respon dari aplikasi tersebut. Perbedaannya pada *load testing* dilakukan pengujian dengan *load condition* tertentu, sedangkan pada *stress testing* dilakukan pengujian pada kondisi ekstrem dengan mensimulasikan sejumlah besar *user* mengakses dan mengirimkan *request* ke sistem, sehingga memberikan beban ke sistem dalam memproses banyak *data* [6] [7]. Dengan melakukan pengujian *load testing* dan *stress testing*, akan diperoleh data terkait dengan kehandalan dan performansi dari aplikasi tersebut, untuk selanjutnya dapat dilakukan analisis agar diperoleh solusi untuk peningkatan aplikasi di masa mendatang [8].

Adapun dalam pengujian *load testing* terdapat metode-metode turunan yang dapat digunakan, dalam penelitian ini penulis menggunakan metode *Black Box Testing*. *Black box testing* akan menguji aplikasi secara fungsional tanpa mengetahui struktur dari program (*no knowledge*) [9]. Sedangkan pada pengujian *stress testing*, penulis menggunakan metode *Gorilla Testing*, dimana modul program diuji berulang kali untuk memastikan ketahanan dan kinerja dari modul tersebut dapat bekerja sesuai dengan fungsinya [9].

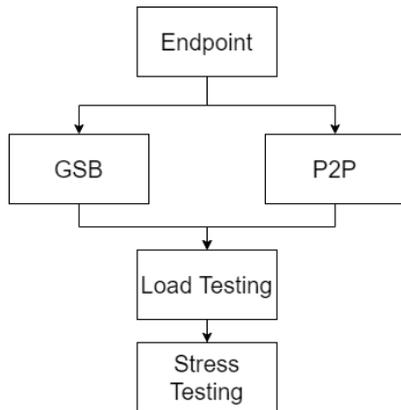
Tools yang penulis gunakan untuk pengujian *load testing* dan *stress testing* adalah Loader.io dan

JMeter. Loader.io adalah sebuah *tool* terbuka yang dikembangkan oleh SendGrid. Loader.io memiliki fitur pengujian *Clients per test*, *Clients per second*, dan *Maintain client load tests* dan menyajikan laporan dalam bentuk *visual* [10]. Paket gratis di Loader.io memungkinkan throughput hingga 10.000 permintaan per detik untuk menjalankan tes beban dalam durasi waktu 1 menit [11]. Apache JMeter adalah sebuah *load and performance testing tool* terbuka yang dibuat menggunakan bahasa pemrograman Java. JMeter bisa digunakan untuk melakukan uji beban pada API dengan *protocol* HTTP. JMeter juga dapat digunakan untuk aplikasi yang menggunakan server/protocol FTP, SMTP, LDAP, Java *Database Connectivity* (JDBC), dan *generic TCP connections* [12].

Penulis mengevaluasi arsitektur interoperabilitas melalui pengujian dengan membandingkan lamanya waktu eksekusi dan tingkat *stress* pada arsitektur interoperabilitas. Perbandingan kecepatan akses dan ketahanan sistem menjadi tolak ukur performa antara arsitektur interoperabilitas GSB dengan P2P. Oleh karena itu, perlu dilakukan pengujian *load test* dan *stress test* untuk mengukur kecepatan akses dan ketahanan sistem. Pengujian *load test* arsitektur interoperabilitas *data* menggunakan GSB dan P2P dapat dilakukan pada *website* PPID LAPAN menggunakan JMeter. Pengujian dilakukan menggunakan kelipatan jumlah pengguna 100, 200, 300, 400, dan 500. Pengguna melakukan akses secara bersamaan dalam 1 detik sehingga dalam 1 detik terdapat 100 user bersamaan melakukan akses dengan menggunakan *protocol REST* dengan *method GET*. Untuk pengujian *stress test*, alat bantu yang digunakan adalah Loader.io. Pada Loader.io dilakukan pengujian 1000 pengguna yang melakukan permintaan secara bersamaan dalam 1 detik selama 1 menit dengan menggunakan *protocol REST* dengan *method GET*.

III. HASIL DAN ANALISIS PENELITIAN

Pada arsitektur interoperabilitas data aplikasi menggunakan GSB dan P2P, integrasi data aplikasi dapat terjadi melalui interkoneksi antara halaman *website* sebagai *REST API client* dengan sumber data sebagai *endpoint REST API*. Pada arsitektur GSB *REST API client* melakukan permintaan data ke *endpoint* melalui GSB, dimana GSB berperan sebagai proxy untuk *endpoint*, sedangkan pada P2P *REST API Client* melakukan permintaan data langsung ke *endpoint* tanpa melalui GSB. Pengujian pada arsitektur GSB, dilakukan dengan mensimulasikan pengguna sebagai *Rest API client* yang mengakses *endpoint* melalui GSB, sedangkan pada arsitektur P2P mensimulasikan pengguna sebagai *Rest API client* mengakses langsung ke *endpoint*.



Gambar 7. Skema Pengujian

Penulis menyiapkan satu *endpoint* sebagai *target data* untuk kedua arsitektur, dimana *endpoint* tersebut digunakan sebagai target pengujian *load testing* dan *stress testing* pada arsitektur GSB maupun P2P. Pengujian dilakukan dengan *tools* JMeter untuk *load testing* dan Loader.io untuk *stress testing*. Penulis membuat dua *thread* pengujian yaitu:

1. *Thread Load Testing*

Pada proses pengujian ini dilakukan simulasi performa dari sisi kecepatan dan ketahanan beban ketika pengguna mengakses *target data*. Pengujian kecepatan akses dilakukan dengan melakukan percobaan dengan kelipatan jumlah pengguna 100, 200, 300, 400, dan 500. Pengguna melakukan akses secara bersamaan dalam 1 detik sehingga dalam 1 detik terdapat 100 user bersamaan melakukan akses dengan menggunakan *protocol REST* dengan *method GET* ke *endpoint* melalui GSB dan P2P. Setiap kelipatan pengujian dilakukan perekaman hasil kecepatan rata-rata, waktu *minimal* akses, waktu maksimal akses, dan *error request*. Pada *thread* ini dilakukan konfigurasi *thread* pada JMeter dengan memasukkan alamat *endpoint* pada kolom *http request*, menetapkan nilai *Number of Threads (users)* dengan kelipatan jumlah pengguna dan *Ramp-up Periods (seconds)* dengan nilai 1 (detik).

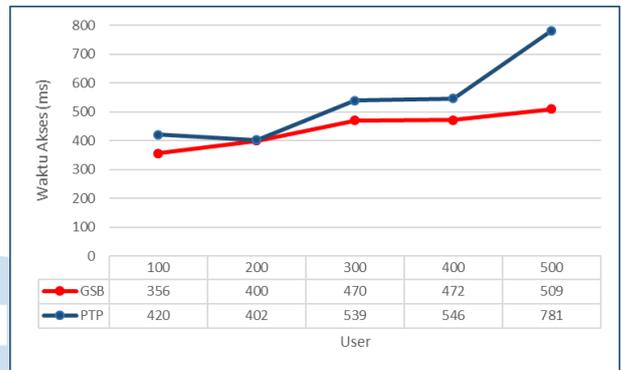
Berdasarkan pengujian dari jumlah *user* 100 hingga 500 dengan kelipatan 100 *user* setiap kali pengujian didapat hasil kecepatan rata-rata akses, waktu minimal akses, waktu maksimal akses, dan *error request* dari setiap percobaan dengan yang dilakukan sebagaimana disajikan pada Tabel 1 dan 2.

Tabel 1. Hasil Pengujian *Load Testing* pada Arsitektur Interoperabilitas GSB

No.	User	Rata-rata (ms)	Waktu Akses (ms)		Error Request (%)
			Minimal	Maksimal	
1	100	356	335	577	0
2	200	400	345	1419	0
3	300	470	339	5063	0
4	400	472	350	1302	0
5	500	509	331	1588	0

Tabel 2. Hasil Pengujian *Load Testing* pada Arsitektur Interoperabilitas P2P

No.	User	Rata-rata (ms)	Waktu Akses (ms)		Error Request (%)
			Minimal	Maksimal	
1	100	420	337	1365	0
2	200	402	332	1889	0
3	300	539	346	2020	0
4	400	546	330	4749	0
5	500	781	428	5675	0



Gambar 8. Perbandingan Hasil Pengujian Interoperabilitas

Berdasarkan perbandingan hasil pengujian interoperabilitas pada Gambar 8, kecepatan akses arsitektur interoperabilitas data aplikasi menggunakan GSB lebih cepat dibanding dengan arsitektur P2P dilihat dari waktu rata-rata akses arsitektur menggunakan GSB lebih kecil dibanding P2P, dengan rata-rata 2ms hingga 309ms. Pada kondisi dengan jumlah beban tertentu (100 sampai 500 user) arsitektur GSB dan P2P dapat melayani *request* dengan tingkat keberhasilan 100%, yang mana menggambarkan kedua arsitektur tersebut memiliki tingkat ketahanan yang sama. Sedangkan dari sisi kecepatan, arsitektur GSB lebih unggul dibanding arsitektur P2P.

2. *Thread Stress Testing*

Pada proses pengujian ini dilakukan simulasi performa dari sisi kecepatan dan ketahanan beban dengan kondisi ekstrem ketika *user* mengakses *target data*. Pengujian *stress testing* dilakukan menggunakan Loader.io dengan akses secara bersamaan sebesar 1000 *user* dalam 1 detik selama 1 menit dengan menggunakan *protocol REST* dengan *method GET* ke *endpoint* melalui GSB dan P2P. Setiap kelipatan pengujian dilakukan perekaman hasil kecepatan rata-rata, waktu minimal akses, dan waktu maksimal akses.

Pada *thread* ini, penulis melakukan konfigurasi dasar pada Loader.io yaitu dengan menambahkan dan verifikasi *target host* *ppid.lapan.go.id*. Selanjutnya penulis menetapkan nilai pada kolom pengujian dengan tipe *Client per seconds*, *Users* sebesar 1000, durasi sebesar 1 menit dan *method GET* pada Loader.io.

Tabel 3. Hasil Perbandingan Pengujian *Stress Testing*

	GSB	P2P
Rata-rata (ms)	11198	14155
Min (ms)	306	229
Max (ms)	55353	83263
Error Request (%)	20,99	4,94
Throughput (request/second)	6,5	5,0

Berdasarkan hasil pengujian *stress testing* arsitektur interoperabilitas data, aplikasi menggunakan GSB lebih cepat dibanding dengan arsitektur P2P dengan selisih waktu rata-rata akses 2957ms, selisih *throughput* sebesar 1.5 request per detik, dan tingkat error pada arsitektur interoperabilitas data menggunakan arsitektur GSB lebih tinggi dari arsitektur P2P dengan selisih tingkat error 16,05%. Pada kondisi ekstrem, arsitektur GSB memiliki tingkat kecepatan dan *throughput* yang lebih tinggi, namun memiliki tingkat *error* yang lebih tinggi dibanding arsitektur P2P.

IV. SIMPULAN

Berdasarkan hasil pengujian *load testing* dan *stress testing*, pada kondisi umum dengan beban tertentu, arsitektur GSB memiliki performansi yang lebih unggul, hal ini terlihat dari waktu akses dan *error request*. Sedangkan pada kondisi ekstrem arsitektur GSB mengalami penurunan performansi, hal ini terlihat dari tingkat *error request* yang lebih besar ketika menangani request dalam jumlah sangat besar.

Pengembangan yang dapat dilakukan pada penelitian selanjutnya adalah melakukan pengujian pada aplikasi lainnya yang menerapkan arsitektur interoperabilitas GSB, P2P, dan arsitektur interoperabilitas alternatif lainnya, yang memiliki keberagaman fitur, kompleksitas, dan dengan memperhatikan perhitungan dan pengalokasian *resource* yang tepat.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Lembaga Penerbangan dan Antariksa Nasional (LAPAN), khususnya satuan kerja Pusat Teknologi

Informasi dan Informasi Penerbangan dan Antariksa yang telah mendukung dan memfasilitasi kegiatan penelitian ini.

DAFTAR PUSTAKA

- [1] NCOIC, "What is Interoperability?," *NCOIC.org*, 2018. [Online]. Available: <http://www.ncoic.org/what-is-interoperability>.
- [2] A. Sasongko, H. A. Aziz, P. Setyantana, D. Sukyadi, and A. Basuki, *MANTRA - Manajemen Integrasi Informasi dan Pertukaran Data (Government Service Bus)*. Jakarta: DIREKTORAT E-GOVERNMENT DIREKTORAT JENDERAL APLIKASI INFORMATIKA KEMENTERIAN KOMUNIKASI DAN INFORMATIKA, 2017.
- [3] A. Haryanto, E. Wiyatnanto, and R. A. Ferianda, "Perancangan Arsitektur Interoperabilitas Data Aplikasi Menggunakan Government Service Bus (GSB)," in *Kajian dan Penerapan Teknologi Informasi dan Komunikasi*, I. G. Prihanto, Suwardi, H. Gunawan, and F. Alusi, Eds. Jawa Tengah: Penerbit Nugra Media, 2018, pp. 43–52.
- [4] D. B. Santoso, A. E. Pramono, and A. G. Persada, "Pengembangan Interoperabilitas Sistem Penanggulangan Gawat Darurat Terpadu (SPGDT) Kabupaten Kebumen," *J. Manaj. Inf. Kesehat. Indones.*, vol. 7, no. 1, p. 43, 2019.
- [5] N. Winastan and R. Ferdiana, "Metode Pengukuran Interoperabilitas untuk Mengukur Kinerja Pelayanan Pemerintah," in *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 2014, pp. 13–18.
- [6] D. Gea, "Pengujian Kualitas Website Ditinjau dari Perspektif Accessibility, Experience, Marketing dan Technology," *ComTech Comput. Math. Eng. Appl.*, vol. 5, no. 1, p. 35, 2014.
- [7] N. M. K. K. S. A. Saha, and D. Chahar, "Comparative Study on Performance Testing with JMeter," *IJARCCCE Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 5, no. 2, pp. 70–76, 2016.
- [8] E. N. Delta and Asmunin, "Performance Test dan Stress Website Menggunakan Open Source Tools," *J. Manaj. Inform.*, vol. 6, no. 1, pp. 208–215, 2016.
- [9] D. I. Permatasari *et al.*, "Pengujian Aplikasi menggunakan Metode Load Testing dengan Apache JMeter pada Sistem Informasi Pertanian," *J. Sist. dan Teknol. Inf.*, vol. 8, no. 1, p. 135, 2020.
- [10] Loader.io, "Test Types," *support.loader.io*, 2015. [Online]. Available: <http://support.loader.io/article/16-test-types>.
- [11] Kamarudin, Kusri, and A. Sunyoto, "Uji Kinerja Sistem Web Service Pembayaran Mahasiswa Menggunakan Apache JMeter (Studi Kasus: Universitas AMIKOM Yogyakarta)," *Teknol. Inf.*, vol. XIII, no. 1, pp. 44–52, 2018.
- [12] T. A. S. Foundation, "Apache JMeter," *jmeter.apache.org*, 2018. [Online]. Available: <http://jmeter.apache.org/>.

Implementasi Algoritma Simon Pada Aplikasi Kamus Perubahan Fi'il (Kata Kerja Bahasa Arab) Berbasis Android

Rahmad Akbar¹, Bambang Pramono², Rizal Adi Saputra³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Halu Oleo, Kendari, 93231, Indonesia
¹akbarbelajar@gmail.com, ²bambang.pramono@uho.ac.id, ³rizaladisaputraa@gmail.com

Diterima 29 November 2020

Disetujui 19 Mei 2021

Abstract— Shorof science or Tashrif is the scientific field of word derivation in Arabic, one focus of the discussion in this field is the process of changing verbs or also known as Fi'il into several other types of words, such as Fi'il Mudhori', Fi'il Madhi, Fi'il Amr, Fi'il Nahi, Isim Fa'il, Isim Maf'ul, Isim Zaman, Isim Makan, Isim Alat, Masdar or Masdar mim. The process of learning Shorof science is still mostly carried out in traditional ways, especially in the pesantren environment by memorizing the derivatives of these words and their translations. While one of the basic books that is often used is the book *Amsilah At-Tashrifiyah* written by KH.Ma'shum bin Ali as a reference for the process of changing words, while looking for a translation in Indonesian must use an Arabic-Indonesian dictionary. This study aims to simplify the word search process by making an android-based dictionary of Fi'il change and utilizing the Simon Algorithm as a word search method, so as to simplify the learning process of Shorof's knowledge. Simon's algorithm is a string matching algorithm where the matching phase is carried out from left to right by initializing each index on a given pattern. After testing, the word search process can be carried out with an average running time of 23.8447333 milli second for searching Arabic words.

Index Terms— String matching algorithm, Simon algorithm, Android Fi'il Dictionary

I. PENDAHULUAN

Bahasa Arab merupakan bahasa resmi di dunia Internasional dan merupakan salah satu bahasa yang kaya akan derivasi atau kata turunan, sebuah kata dalam Bahasa Arab terkadang bisa dikembangkan menjadi puluhan bahkan ratusan. Bidang keilmuan derivasi dalam Bahasa Arab disebut juga dengan ilmu *Shorof* atau *Tashrif*, sedangkan kata dalam Bahasa Arab terbagi menjadi tiga kelompok yaitu *Isim* (kata benda), *Fi'il* (kata kerja) dan *Khurf* (huruf yang memiliki arti) [1]. Diantara ketiga kelompok kata tersebut *Fi'il* menjadi kelompok kata yang memiliki banyak turunan dan perubahan, sehingga perlu adanya pembahasan tersendiri seperti yang terangkum dalam kitab *Tashrif* yang berjudul *Amsilah At-Tashrifiyah* yang ditulis oleh KH.Ma'shum bin Ali, buku tersebut kusus membahas tentang perubahan-perubahan *Fi'il*,

antara lain *Fi'il Mudhore'* (kata kerja bentuk sekarang dan yang akan datang), *Fi'il Madhi* (kata kerja lampau), *Masdar* (kata kerja yang dibendakan) dan masih banyak lagi. Akan tetapi daftar kata pada kitab *Amsilah At-Tashrifiyah* tersebut tidak memiliki terjemahan secara langsung dalam bahasa Indonesia, sehingga pengguna harus memiliki kamus bahasa Arab-Indonesia untuk mengetahui arti dari daftar kata tersebut beserta turunan-turunannya.

Berdasarkan beberapa uraian di atas, menegaskan bahwa perlunya sebuah aplikasi yang dapat membantu memudahkan proses pembelajaran Bahasa Arab, dan yang lebih penting lagi dapat memudahkan proses pencarian *Fi'il* beserta turunan-turunan dan terjemahannya dalam bahasa Indonesia. Dengan adanya solusi pembuatan tersebut, dibutuhkan sebuah metode yang dapat digunakan dalam proses pencarian kata pada aplikasi yang akan dibangun, salah satu metode yang dapat digunakan antara lain ialah Algoritma *Simon*.

Algoritma *Simon* merupakan sebuah algoritma pencocokan *string* dengan fase pencariannya dilakukan dari kiri ke kanan dengan tahapan inialisasi tiap indeks pada pola yang diberikan [9]. Berdasarkan beberapa penelitian Algoritma *Simon* terdahulu, yaitu penelitian yang telah dilakukan oleh Putri Chaliska pada tahun 2017 dengan membandingkan Algoritma *Simon* dengan Algoritma *On Ordered Alphabets* dan memperoleh kesimpulan bahwa Algoritma *Simon* membutuhkan waktu yang lebih cepat dengan rata-rata waktu 1,5134 ms dan kompleksitas algoritma yang diperoleh $\Theta(n)$, sedangkan Algoritma *On Ordered Alphabets* memiliki rata-rata waktu 1,6928 ms [3]. Kemudian Megawaty pada tahun 2018 juga melakukan penelitian dengan membandingkan Algoritma *Simon* dengan Algoritma *Brute Force* dan memperoleh kesimpulan bahwa Algoritma *Simon* membutuhkan waktu lebih cepat dengan rata-rata waktu 1,8 ms, sedangkan Algoritma *Brute Force* membutuhkan rata-rata waktu 5,5 ms [9]. Adapun *string* yang digunakan pada kedua penelitian tersebut berupa *string* dalam Bahasa Indonesia, sedangkan

pada penelitian ini *string* yang digunakan dalam Bahasa Arab.

II. LANDASAN TEORI

A. Ilmu Shorof

Ilmu *Shorof* adalah ilmu yang membahas tentang derivasi atau perubahan kata dalam Bahasa Arab. Perlu diketahui mempelajari Bahasa Arab termasuk didalamnya ilmu Nahwu dan ilmu *Shorof* hukumnya wajib. Karena mempelajari keduanya adalah merupakan sarana untuk dapat memahami serta mendalami makna yang terkandung dalam Al-Qur'an.

Ilmu *Shorof* juga dinamakan dengan *Ummul'ulum* (induknya ilmu) karena dari ilmu *Shorof* itu kita dapat mengetahui berbagai macam bentuk perubahan pecahan-pecahan kata yang antara kata satu dengan kata yang lainnya mempunyai arti yang berbeda. Sebagai contoh lafadz *fatahan* = فَتْحًا mempunyai arti pembukaan, *fatihun* = فَاتِحٌ mempunyai arti orang yang membuka, *miftāhun* = مِفْتَاحٌ mempunyai arti alat untuk membuka. semua ini berasal dari satu kata yaitu *fataha* = فَتَحَ yang artinya membuka.

Menurut arti bahasa (*lugat*) *Shorof/Tasrif* yaitu berubah atau mengubah dari bentuk aslinya kepada bentuk yang lain. Misalnya merubah bentuk bangunan rumah kuno menjadi bentuk bangunan rumah modern. Sedangkan menurut istilah (menurut kalangan ulama *Shorof*), yaitu berubahnya bentuk asal pertama *Fi'il Maḍhi* menjadi *Fi'il Muḍhori'*, *Fi'il Amr*, *Fi'il Nahi*, *Isim Zaman*, *Isim Makan*, dan terakhir sampai pada *Isim Alat* [7].

B. Algoritma String Matching

Pencocokan *string* (*string matching*) adalah proses pencarian kata (*Pattern*) dalam sebuah teks. *String matching* bisa digunakan untuk menemukan satu *Pattern* yang pertama kali ditemukan, atau yang lebih umum menampilkan semua *Pattern* yang dapat ditemukan dalam teks. Jenis *string matching* bermacam-macam, dibedakan menurut hasil yang diinginkan [4]. Prinsip kerja algoritma *string matching* adalah sebagai berikut :

- Memindai teks dengan bantuan sebuah *window* yang ukurannya sama dengan *Pattern*.
- Menempatkan *window* pada awal teks
- Membandingkan karakter pada *window* dengan karakter dari *Pattern*.

Setelah pencocokan (baik hasilnya cocok atau tidak cocok), dilakukan *shift* ke kanan pada *window*. Prosedur ini dilakukan berulang-ulang sampai *window* berada pada akhir teks. Mekanisme ini disebut *sliding-window*.

String matching dibagi menjadi dua, yaitu *exact string matching* dan *heuristic (statistical matching)*.

Exact matching digunakan untuk menemukan *Pattern* yang berasal dari satu teks. Contoh pencarian *exact matching* adalah pencarian kata “pelajar” dalam kalimat “saya seorang pelajar” atau “saya seorang siswa”. Sistem akan memberikan hasil bahwa kata pelajar dan siswa bersinonim. Algoritma *exact matching* diklasifikasikan menjadi tiga bagian menurut arah pencariannya, yaitu :

- *Arah pencarian dari kiri ke kanan.*
Algoritma yang termasuk dalam kategori ini adalah *Brute Force*, *Simon*, *Morris* dan *Pratt* (yang kemudian dikembangkan oleh *Knuth*, *Morris* dan *Pratt*).
- *Arah pembacaan dari kanan ke kiri.*
Algoritma yang termasuk dalam kategori ini adalah *Boyer Moore* yang kemudian dikembangkan menjadi algoritma *Turbo Boyer Moore*, *Tuned Boyer Moore* dan *Zhu Takaoka*.
- *Arah pencarian yang ditentukan oleh program.*
Algoritma yang termasuk dalam kategori ini adalah algoritma *Colussi* dan *Crochemore-Perrin*. *Heuristic matching* adalah teknik yang digunakan untuk menghubungkan dua data terpisah ketika *exact matching* tidak mampu mengatasi karena pembatasan pada data yang tersedia. *Heuristic matching* dapat dilakukan dengan perhitungan *distance* antara *Pattern* dengan teks. *Exact* dan *heuristic matching* memiliki kemiripan makna tetapi berbeda tulisan.

C. Algoritma Simon

Pada Algoritma *Simon* akan dilakukan fase untuk pencocokan *string* yang dilakukan dari kiri ke kanan dengan tahapan inisialisasi di setiap indeksnya berdasarkan *Pattern* yang akan diberikan, kemudian pada keadaan awal pencarian *string* diberi nilai -1 yang artinya *mismatch* atau tidak sesuai dengan *Pattern*. Berikut adalah contoh Algoritma Simon beserta simulasinya dalam melakukan pencocokan *Pattern* terhadap teks [3].

- *User* akan mencari kata “مَنْصُورٌ” dengan memasukan *Pattern* “صُورٌ”.
- Kata dan *Pattern* akan dibalik menggunakan fungsi *reverse* sehingga menjadi “رُوصْنَمٌ” dan “رُوصُنٌ”.
- Setelah *Pattern* melalui proses pembalikan, maka sistem langsung memproses dengan mendeklarasikan terlebih dahulu yaitu jumlah array *Pattern* $m = \text{Pattern.length}()$, jumlah array setiap kata yang ada di *database* $n = \text{text.length}()$, $\text{Result} = \text{False}$, $j = 0$ dan inisialisasi $x = -1$.
- Kemudian sistem akan melakukan proses pencocokan dengan syarat yang pertama yaitu $j \leq n-1$.
- Selanjutnya syarat yang kedua ialah $x+1 < m$.

- Setelah dua syarat diatas terpenuhi, sistem langsung melakukan proses pencocokkan karakter dimulai dari pencocokkan karakter array ke $x+1$ dari *Pattern* dengan karakter array ke j dari kata yang ada pada *database* atau ditulis dengan $Pattern(x+1)=text(j)$.
- Apabila terjadi kecocokkan maka nilai x akan ditambah 1, namun apabila terjadi ketidakcocokkan maka x diberi nilai -1, sedangkan nilai j selalu ditambah 1.
- Proses akan terus berulang hingga nilai $x = m-1$ dengan kata lain semua array *Pattern* telah dicocokkan.

Inisialisasi dimulai dari -1.

- Keadaan : -1 {ص → 0}
 - Keadaan : 0 {ص → 0, ُ → 1}
 - Keadaan : 1 {ُ → 1, و → 2}
 - Keadaan : 2 {و → 2, ُ → 3}
 - Keadaan : 3 {ُ → 3, ر → 4}
 - Keadaan : 4 {ر → 4, ُ → 5}
 - Keadaan : 5 { }
- Berhenti di keadaan 5.

Tabel 1. Pencocokan Pertama

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ
P	ص									

-1

Karna *Pattern* tidak sesuai dengan teks maka م diberi nilai -1 yang berarti *mismatch*.

Tabel 2. Pencocokan Kedua

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ
P		ص								

-1

Karna *Pattern* tidak sesuai dengan teks maka َ diberi nilai -1 yang berarti *mismatch*.

Tabel 3. Pencocokan Ketiga

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ
P			ص							

-1

Karna *Pattern* tidak sesuai dengan teks maka ن diberi nilai -1 yang berarti *mismatch*.

Tabel 4. Pencocokan Keempat

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ	م
P				ص							

-1

Karna *Pattern* tidak sesuai dengan teks maka ُ diberi nilai -1 yang berarti *mismatch*.

Tabel 5. Pencocokan Kelima

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ
P					ص					

0

Karna *Pattern* sesuai dengan teks maka ص diberi nilai 0 yang artinya *match*.

Tabel 6. Pencocokan Keenam

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ
P						ُ				

1

Karna *Pattern* sesuai dengan teks maka ُ diberi nilai 1 yang artinya *match*.

Tabel 7. Pencocokan Ketujuh

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ
P							و			

2

Karna *Pattern* sesuai dengan teks maka و diberi nilai 2 yang artinya *match*.

Tabel 8. Pencocokan Kedelapan

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ
P								ُ		

3

Karna *Pattern* sesuai dengan teks maka ُ diberi nilai 3 yang artinya *match*.

Tabel 9. Pencocokan Kesembilan

T	م	َ	ن	ُ	ص	ُ	و	ُ	ر	ُ
P									ر	

4

Karna *Pattern* sesuai dengan teks maka ر diberi nilai 4 yang artinya *match*.

Tabel 10. Pencocokan Kesepuluh

T	م	ن	ص	و	ر	5
P					ر	5

Karna *Pattern* sesuai dengan teks maka 5 diberi nilai 5 yang artinya *match*. Keadaan 5 telah terpenuhi, sesuai inisialisasi awal pencarian akan berhenti pada keadaan 5.

D. JSON(Java Script Object Notation)

JSON(Java Script Object Notation) adalah format pertukaran data yang ringan, mudah dibaca dan ditulis oleh manusia, serta mudah diterjemahkan dan dibuat(generate) oleh komputer. Format ini dibuat berdasarkan dari bahasa pemrograman Java Script, Struktur data ini disebut sebagai struktur data universal. Pada dasarnya semua bahasa pemrograman modern mendukung struktur dalam bentuk yang sama maupun berlainan. Hal ini pantas disebut demikian karena format data mudah dipertukarkan dengan bahasa-bahasa pemrograman yang juga berdasar pada stuktur data ini. *JSON* menggunakan bentuk sebagai berikut.

Objek adalah nama/nilai yang tidak terurutkan. Objek dimulai dengan "{" (kurung kurawal buka) dan diakhiri dengan "}" (kurung kurawal tutup). Setiapnama diikuti dengan ":" (titik dua) dan setiap pasangan nama/nilai dipisahkan oleh "," (koma)[6]

E. Case Folding

Case Folding adalah sebuah proses penyamaan besar kecilnya huruf pada setiap kata serta pembuangan tanda baca yang masih terdapat pada suatu kalimat [8]. Berikut adalah beberapa proses *Case Folding*

- Ubah kata menjadi huruf kecil
Pada proses ini semua string yang terdapat pada *Pattern* akan diubah menjadi huruf kecil menggunakan fungsi *toLowerCase()*.
- Buang tanda baca
Pada tahap ini hasil dari proses penyamaan huruf menjadi huruf kecil diolah kembali untuk menghilangkan *noise* berupa tanda baca dan angka yang masih terdapat pada huruf. Pada proses ini digunakan fungsi *replace()*, delimiter yang digunakan adalah *\d* yang menandakan digit[0-9] dan *\W* yang menandakan simbol atau karakter yang tidak termasuk kedalam huruf dan pengganti merupakan karakter yang digunakan untuk mengganti karakter yang sesuai dengan delimiter, pada kasus ini digunakan spasi sebagai karakter pengganti. Dengan delimiter tersebut fungsi akan menghapus angka dan simbol yang terdapat pada kata yang dimasukan sebagai parameter.

III. IMPLEMENTASI DAN PEMBAHASAN

A. Implementasi Rancangan Antarmuka User

Implementasi *interface* dari perangkat lunak dilakukan berdasarkan rancangan yang telah dibuat pada bab sebelumnya, namun disesuaikan dengan komponen-komponen yang tersedia pada Flutter. Implementasi antarmuka ditampilkan dalam bentuk screenshot aplikasi Android.

- *Beranda*

Beranda merupakan tampilan awal ketika aplikasi pada aplikasi Kamus perubahan Fi'il dan berisi penjelasan singkat tentang aplikasi.



Gambar 1. Tampilan Menu Beranda User

- *Menu Istilah*

Menu istilah menampilkan daftar istilah yang digunakan pada aplikasi kamus perubahan *Fi'il*.



Gambar 2. Tampilan Menu Istilah

- *Menu Wazan*

Menu Wazan menampilkan daftar wazan yang terdapat pada aplikasi Kamus perubahan *Fi'il* beserta daftar kata dan turunan-turunannya dengan cara sebagai berikut:

- Pilih menu daftar *Wazan*, maka daftar *Wazan* akan ditampilkan.

- Pilih salah satu *Wazan*, maka daftar kata yang terdapat pada *Wazan* tersebut akan ditampilkan dalam bentuk *Fi'il Madhi*.
- Pilih salah satu kata dari daftar kata tersebut, maka akan ditampilkan daftar turunan beserta terjemahan dan *audio* dari kata tersebut.



Gambar 3. Tampilan Menu *Wazan*

- **Menu Pencarian**
 Menu Pencarian berisi *form* yang dapat digunakan mencari kata yang terdapat pada Kamus perubahan dengan *cra* sebagai berikut:
 - Pilih menu pencarian.
 - Pilih salah satu kategori pencarian yang menggunakan Bahasa Indonesia atau Bahasa Arab.
 - Masukkan kata yang akan dicari pada *form* pencarian.
 - Apabila kata ditemukan, maka akan ditampilkan daftar kata tersebut dibawah *form*, baik itu berupa kata dalam bentuk *Fi'il Mudhori*, *Fi'il Madhi*, *Fi'il Amr*, *Fi'il Nahi*, *Isim Fa'il*, *Isim Maf'ul*, *Isim Zaman*, *Isim Makan*, *Isim Alat*, *Masdar* maupun *Masdar mim*.
 - Pilih salah satu kata yang ditemukan, maka akan ditampilkan daftar turunan dari kata tersebut.



Gambar 4. Tampilan Menu Pencarian

- **Menu Tentang**
 Menu tentang berisi informasi mengenai deskripsi ataupun spesifikasi dari sistem aplikasi Kamus Perubahan *Fi'il*.



Gambar 5. Tampilan Menu Tentang

B. Implementasi Coding Algoritma Simon

Dalam proses coding, dibuat dua algoritma Simon yang masing-masing digunakan dalam proses pencocokan *string* Bahasa Arab. Adapun implementasinya adalah sebagai berikut :

- Deklarasi *String text* sebagai variabel *baliktext* yang berasal dari variabel *arb*, yakni *string* kata yang dipanggil dari data *JSON* dan variabel *arb* tersebut diproses dengan fungsi *runes.toList.reversed()* untuk membalik kata tersebut.
- Deklarasi *String pattern* sebagai *balikpattern* yang berasal dari variabel *_filter*, yakni kata yang dimasukan oleh *user* dan variabel *_filter* tersebut diproses dengan fungsi *runes.toList.reversed()* untuk membalik kata tersebut.
- Deklarasi *int n = text.length*, sebagai panjang *string* dari variabel *text*, kemudian *int m = pattern.length*, sebagai panjang variabel *pattern*.
- Deklarasi *int j=0;* sebagai awal dari proses pencocokan *string* yang dimulai dari sebelah kiri atau *array* ke-0;
- Deklarasi *int x=-1;* sebagai inialisasi awal proses pencocokan *string*.
- Deklarasi *result = false;* sebagai keadaan awal proses dan ketika terjadi ketidak cocokan pada hasil pencocokan *string*.
- Kemudian akan dilakukan perulangan jika kondisi $j \leq n-1$ dan j selalu ditambah satu.
- Kemudian dalam perulangan terdapat dua kondisi, yaitu kondisi pertama $x+1 \leq m-1$ dan kondisi kedua $x == m-1$.
- Dalam kondisi $x+1 \leq m-1$ terdapat sebuah kondisi jika benar $pattern[x+1] == text[j]$ maka $x+1$, dan jika tidak maka $x=-1$.

- *Pattern* dinyatakan *match* dengan kata yang dicari apabila kondisi $x == m-1$ sama dengan *true*, sehingga *result = true* yang berfungsi untuk menampilkan hasil pencarian dan variabel *jkata++* untuk menghitung jumlah kata yang ditemukan, serta $j=n-1$ untuk mengakhiri proses.

Gambar 6. Coding Algoritma Simon

C. Pengujian Algoritma Simon dalam Proses Pencarian Kata

Pengujian ini dilakukan untuk menguji kecepatan sistem dalam melakukan proses pencarian sebuah kata berdasarkan *Pattern* yang dimasukan. Pengujian ini dibagi dalam dua hasil proesees, antarlain proses pencarian kata dengan *Pattern* dalam Bahasa Arab.

- Pencarian satu sample kata dalam beberapa *Pattern*

Tabel 11. Pengujian Pencarian Satu Kata

No	Pattern	Panjang Pattern	Waktu pencarian Ms (milisecond)
1	ن	1	17,186
2	نَ	2	19,33
3	نص	3	14,697
4	نصَ	4	16,356
5	نصر	5	21,104
6	نصرَ	6	13,27
Rata-rata waktu running time			16,9905

- Pencarian 10 sample kata yang berbeda

Tabel 12. Pengujian Pencarian 10 Kata

No	Pattern	Panjang Pattern	Waktu pencarian Ms (milisecond)
1	مَدَ	4	25,477
2	صَانَ	5	38,474
3	غَزَا	5	22,331
4	أَمَلَ	6	14,357
5	ضَرَبَ	6	26,349
6	فَرَ	4	28,512
7	وَعَدَ	6	21,013
8	يَسَرَ	6	26,392
9	سَارَ	5	22,605
10	شَوَى	5	20,873
Rata-rata waktu running time			24,638

- Pencarian 10 sample kata yang tidak ditemukan

Tabel 13. Pengujian Pencarian 10 Kata

No	Pattern	Panjang Pattern	Waktu pencarian Ms (milisecond)
1	يَفَعُ	6	37,784
2	وَجِنَ	6	20,928
3	بَيْسَ	6	30,212
4	حَصِنَ	6	28,014
5	جِبِنَ	6	31,795
6	يَمِنَ	6	34,868
7	سَرُوَ	6	19,043
8	أَدَبَ	6	30,344
9	لَوْمَ	6	28,156
10	نَخْرَجَ	8	37,913
Rata-rata waktu running time			29,9057

- Rata-rata yang diperoleh dari pengujian 1-3 string Bahasa Arab

Tabel 14. Rata-rata running time pencocokan string

Jenis Pengujian	Running Time ms(mili second)
Pengujian 1	16,9905
Pengujian 2	24,638
Pengujian 3	29,9057
Rata-rata	23,8447333

IV. SIMPULAN

Berdasarkan analisis hasil pengujian maka dapat disimpulkan bahwa Algoritma Simon dapat diimplementasikan sebagai metode pencocokan *string* pada proses pencarian kata dalam kamus perubahan *Fi'il*, baik itu berupa kata dalam Bahasa Arab maupun Bahasa Indonesia. Adapun rata-rata *running time* yang dibutuhkan oleh Algoritma Simon dalam proses pencocokan *string* dalam penelitian ini adalah 23,8447333 *mili second*, sedangkan proses *running time* tercepat yang diperoleh proses *running time* tercepat yang diperoleh 13,27 *mili second* dan *running time* terlama 38,474 *mili second*.

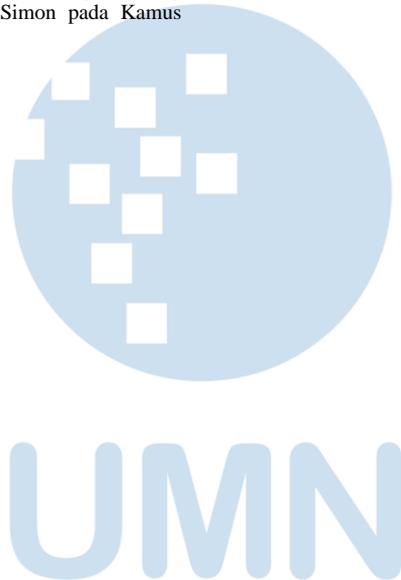
Adapun saran untuk penelitian selanjutnya mengenai "Implementasi Algoritma Simon Pada

Aplikasi Kamus Perubahan *Fi'il* (Kata Kerja Bahasa Arab) Berbasis *Android* yaitu :

1. Sistem aplikasi Kamus Perubahan *Fi'il* ini masih berbasis *offline*, sehingga ukuran aplikasi cukup besar, sehingga kurang efektif apabila ditambah lebih banyak lagi kata dan *audio* yang dimasukan.
2. Sistem penyimpanan data pada aplikasi ini hanya menggunakan data *JSON* dan tidak menggunakan *database*, sehingga proses *update* maupun *delete* dilakukan secara manual dalam program aplikasi, sehingga harus dilakukan *build* aplikasi kembali.

DAFTAR PUSTAKA

- [1] Alkhatib, A. L. bin M. (2016). Al Khathib Ensiklopedia Komplit Menguasai Shorof Tashrif (1st ed.). Mitra Pustaka. pustakapelajar@yahoo.com
- [2] Andriani, A. (2015). Urgensi Pembelajaran Bahasa Arab dalam Pendidikan Islam. Ta'allum: Jurnal Pendidikan Islam, 3(1), 39–56. <https://doi.org/10.21274/taalum.2015.3.1.39-56>.
- [3] Chaliska, P. (2017). Perbandingan Algoritma String Matching On Ordered Alphabet dan Algoritma Simon pada Kamus Istilah Arsitektur. USU E-JOURNAL. <http://repositori.usu.ac.id/handle/123456789/4112>.
- [4] Charras, C. and, & Lecroq, T. (2004). Handbook of Exact String Matching Algorithms. 238.
- [5] Community, Ewolf. 2011. Buku Wajib Programmer: Indeks Lengkap Syntax Kumpulan Perintah-perintah Dasar Pemrograman yang Sering Digunakan Visual Basic, Delphi, C++, HTML, JavaScript, PHP, dan SQL. Yogyakarta: MediaKom.
- [6] Eka Rinjani, R. (2012). Pemanfaatan Json (Javascript Object Notation) Sebagai Data Interchange Pada Sistem Automatic Testing Dan Web Based Learning D3 Teknik Informatika Uns Tugas.
- [7] Kandir, N. (2017). Ringkasan Al-Amtsilah At-Tashrifiyah (1st ed.). Pustaka Syabab.
- [8] Langgeni, Baizal & Firdaus. (2010). Clustering Artikel Berita Berbahasa Indonesia Menggunakan Unsupervised Feature Selection, Seminar Nasional Informatika, Yogyakarta.
- [9] Megawaty. (2017). Implementasi dan Perbandingan Algoritma Brute Force dengan Algoritma Simon dalam Pembuatan Kamus Istilah Kebidanan. Universitas Sumatera Utara. <http://repositori.usu.ac.id/handle/123456789/4660>
- [10] Parno, dharmayanti, N. (2011). Aplikasi Mobile Kamus Istilah_Ug.Pdf. <http://repository.gunadarma.ac.id>



Implementasi Algoritma Complement dan Multinomial Naïve Bayes Classifier Pada Klasifikasi Kategori Berita Media Online

Muhammad Naufal Randhika¹, Julio Christian Young², Alethea Suryadibrata³, Hadian Mandala⁴

^{1,2,3}Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia

⁴Faculty of Engineering, Universitas Hamzanwadi, Nusa Tenggara Barat, Indonesia

¹muhammad.randhika@student.umn.ac.id, ²julio.christian@umn.ac.id, ³alethea@umn.ac.id,

⁴hadian_mandala@hamzanwadi.ac.id

Diterima 13 Januari 2021

Disetujui 23 Mei 2021

Abstract— The development of computer technology and the dissemination of information via internet are increasing significantly from time to time. News is one of the information media which also increased. Conventional printed media has now been replaced by electronic media known as online news portals/ media. PT Merah Putih Media is one of developing online news media. There are three main categories (Lifestyles, Sports, and Indonesia) in its portal that still categorized manually by the editor in chief within the company. This study tried to test the suitability of two classification algorithms that able to replace the manual process, namely Multinomial Naïve Bayes (MNBC) and Complement Naïve Bayes (CNBC). Moreover, experiments related to the combination of both of algorithms was also tried. Based on series of experiments which had conducted, we found that the combination of MNBC and CNBC models are able to achieve the F1-Score of 90.13%.

Keywords— PT Merah Putih Media, Text Mining, Multinomial Naïve Bayes Classifier, Complement Naïve Bayes Classifier

I. LATAR BELAKANG

Perkembangan teknologi dan penyebaran informasi di internet selalu meningkat dari waktu ke waktu. Berita merupakan salah satu media informasi yang juga turut mengalami peningkatan [1].

PT. Merah Putih Media merupakan media berita online. Media berita online milik PT. Merah Putih dapat diakses pada alamat URL merahputih.com. Berita yang disampaikan terdiri dari tiga kategori mulai dari berita tentang Indonesia, Hiburan dan Gaya Hidup, serta Olahraga. Namun, pembagian artikel berita ke dalam kategori dilakukan secara manual oleh kepala redaksi jurnalis. Padahal hal ini bukan bagian dari deskripsi pekerjaannya dan hingga saat ini belum ada yang melamar pekerjaan operator kategori berita. Hal ini sering merepotkan ketika jumlah berita yang ingin diterbitkan banyak. Maka dari itu dibutuhkan sebuah sistem untuk mengklasifikasi berita [2].

Penelitian terhadap klasifikasi kategori berita merahputih.com milik PT Merah Putih Media akan menggunakan teknik Text Classification dengan algoritma Multinomial Naïve Bayes Classifier dikombinasikan dengan algoritma Complement Naïve Bayes Classifier dan menggunakan ekstraksi fitur TF-IDF.

Penelitian sebelumnya yang berkaitan dengan klasifikasi berita menggunakan algoritma Multinomial Naïve Bayes Classifier pernah dilakukan oleh Rahman, Wiranto, dan Afrizal, di mana penelitiannya mencoba mengklasifikasikan berita yang diperoleh berdasar lima media *online* ternama yaitu Detik, Viva, inilah.com, AntaraNews dan Okezone. Penelitian ini mencoba mengkategorikan berita-berita tersebut ke dalam tiga kategori besar yaitu kategori Kesejahteraan Rakyat, Ekonomi, dan Politik, Hukum dan Keamanan (Polhukam). Hasil penelitian sebelumnya menunjukkan bahwa metode Multinomial Naïve Bayes Classifier dan metode ekstraksi fitur TF-IDF menghasilkan nilai akurasi akhir sebesar 94,29% [3].

Penelitian lain yang berkaitan dengan klasifikasi berita dilakukan oleh Siti Nur Asiyah di mana penelitiannya membandingkan metode Support Vector Machine (SVM) dan K-Nearest Neighbor (KNN) untuk mengklasifikasi berita online. Penelitian ini menggunakan data yang berasal dari situs berita online detik.com. Penelitian ini mencoba mengklasifikasikan berita-berita ke dalam lima kategori yang terdiri dari news, finance, hot, sport, dan otomotif. Hasil dari penelitian menunjukkan bahwa metode SVM dapat menghasilkan nilai akurasi sebesar 93,14% dan metode K-NN menghasilkan nilai akurasi sebesar 68,90% [4].

Penelitian yang serupa tentang klasifikasi berita Online juga dilakukan oleh Bening, Dian, dan Lailil, di mana penelitiannya bertujuan untuk menghindari *human error* ketika akan menerbitkan berita di kategori yang tidak tepat pada portal berita online. Penelitian ini mencoba menguji tingkat performa dari metode pembobotan TF-IDF dan metode Cosine Similarity. Hasil dari penelitian menunjukkan bahwa gabungan

dari kedua metode mampu memperoleh nilai akurasi sebesar 91,25% [5].

Berdasarkan kajian penelitian di atas, dapat dilihat bahwa algoritma Multinomial Naïve Bayes memiliki akurasi yang paling baik. Maka penelitian ini juga menggunakan algoritma Multinomial Naïve Bayes dan untuk meningkatkan akurasi klasifikasi, penelitian ini akan mengkombinasikan dengan algoritma Complement Naïve Bayes. Menurut [6] bahwa algoritma ini dapat memvalidasi prediksi klasifikasi dari algoritma Multinomial Naïve Bayes.

II. STUDI PUSTAKA

A. Text Mining

Text Mining adalah serangkaian metode komputasi yang memiliki kemampuan untuk menggali informasi dalam kumpulan data teks dan menjalankan tugas klasifikasi untuk sebuah dokumen masukan secara otomatis. Sebagai sebuah bidang ilmu, *Text Mining* merupakan sub-bidang dari keilmuan *Data Mining* yang memiliki kapabilitas untuk mengekstraksi pola tersembunyi dari sekumpulan data tekstual dalam jumlah yang masif. Istilah *Text Classification* merupakan salah satu bentuk pendekatan *Text Mining* yang dapat digunakan untuk mengembangkan sebuah model klasifikasi teks [7].

Text Classification merupakan pemberian label dengan kategori yang sudah ditentukan pada sebuah teks dengan bahasa alami [8]. Klasifikasi teks diterapkan dalam berbagai konteks, mulai dari pengindeksan sebuah dokumen berdasarkan kosa kata, pemfilteran sebuah dokumen, pembuatan metadata otomatis, dan pada macam aplikasi lainnya [8]. Ada beberapa cara umum dalam klasifikasi teks secara otomatis, yaitu *pre-processing*, *feature extraction/selection*, *modeling* menggunakan teknik pembelajaran mesin, serta *training* dan *testing* pada *classifier* [9].

Text Processing merupakan salah satu tahapan dalam *Text Mining*. *Text Processing* diimplementasikan untuk mentransformasikan data tekstual yang tidak terstruktur sehingga menjadi lebih terstruktur sebelum dilakukan tugas-tugas *Text Mining* lainnya [10].

B. Text Preprocessing

Tahap *Text Preprocessing* tersusun atas tahap-tahap lain yang lebih kecil, yaitu *Case Folding*, *Tokenisasi*, *Filtering*, dan *Stemming*. Tahap-tahap ini bersifat opsional dan bergantung pada kebutuhan analisis data yang akan dilakukan [11].

Case Folding merupakan salah satu bentuk teknik *text preprocessing*. Tujuan dari proses ini adalah untuk menormalisasikan data dengan mengubah seluruh huruf yang muncul dalam kumpulan dokumen menjadi huruf kecil [3]. Pada proses ini juga dilakukan tahapan untuk menghilangkan karakter-karakter yang dianggap tidak relevan dalam sebuah tugas analisis teks seperti: penghilangan tanda baca, angka dan karakter lain

selain huruf alfabet [11]. Kemudian, pada tahap ini juga akan dilakukan penghapusan jumlah spasi yang berlebihan di awal dan akhir sebuah kata, teknik ini biasa disebut *whitespace removal* [11].

Tokenisasi merupakan proses pencacahan sebuah *string* masukan menjadi unit-unit terkecil (kata) yang menyusunnya [11]. Pada prinsipnya, tujuan dari dilakukan proses ini adalah untuk mengetahui unit-unit terkecil yang menyusun sebuah dokumen [12].

Filtering merupakan proses pemilihan kata-kata yang dianggap relevan untuk sebuah tugas analisis teks dari hasil tokenisasi. Proses *filtering* juga dikenal dengan istilah lain yakni, *stopword removal*. Dalam proses *stopword removal*, setiap kata yang muncul dalam dokumen akan diperiksa apakah kata tersebut menjadi bagian dari *stop list* ataupun *word list*. *Stop list* merujuk pada kumpulan kata yang dianggap tidak penting dan tidak memiliki makna tertentu untuk sebuah tugas analisis, sedangkan *word list* merujuk pada kumpulan kata yang memiliki definisi sebaliknya. Saat sebuah kata tergabung ke dalam *stop list*, kata tersebut akan dihilangkan dari teks aslinya [11].

Stemming merupakan proses yang dijalankan untuk mengembalikan sebuah kata menjadi kata dasar atau dengan kata lain proses menemukan kembali akar kata dari setiap kata yang menyusun sebuah dokumen terfilter. Setelah proses *stemming* selesai dijalankan, setiap kata yang berimbuhan akan kembali ke bentuk kata dasarnya. Proses *stemming* bertujuan untuk melakukan normalisasi terhadap data dan menghindari representasi teks dengan distribusi kata yang tersebar (*sparse distributed representations*) [11].

C. TF-IDF

Metode TF-IDF adalah salah satu metode ekstraksi fitur yang bekerja dengan merepresentasikan sebuah dokumen menjadi sekumpulan bobot tertentu. Dalam metode TF-IDF, sekumpulan bobot yang merepresentasikan sebuah dokumen dihitung berdasarkan kemunculan setiap kata dalam sebuah dokumen yang telah dinormalisasikan dengan frekuensi kemunculan kata tersebut pada kumpulan dokumen. Proses normalisasi pada rumus perhitungan bobot TF-IDF bertujuan untuk mengetahui derajat kepentingan dari setiap kata yang menyusun sebuah dokumen. Berdasarkan penjelasan yang telah diberikan, rumus pembobotan tiap kata (*term - t*) dalam sebuah dokumen *d* dapat dilihat pada bagian di bawah ini [13],

$$w_{t,d} = tf_{t,d} \times \log\left(\frac{N}{df_t}\right) \quad (1)$$

dengan,

$w_{t,d}$ = Bobot TF-IDF

$tf_{t,d}$ = Banyaknya kata *t* dalam sebuah dokumen *d*

N = Frekuensi dokumen

df_t = Frekuensi dokumen yang mengandung kata t

D. Multinomial dan Gaussian Naïve Bayes Classifier

Multinomial Naïve Bayes Classifier (MNBC) merupakan model yang dibentuk berdasarkan Teorema Bayes yang terbukti memiliki performa yang baik untuk tugas klasifikasi teks. Dalam persamaan MNBC, tidak seperti persamaan pada algoritma Bernoulli Naïve Bayes yang hanya memperhitungkan jumlah kemunculan dokumen dengan kata tertentu dalam suatu kelas, frekuensi kemunculan kata dalam suatu kelas juga turut diperhitungkan [14]. Definisi formal dari MNBC untuk mengklasifikasikan kelas c untuk sebuah dokumen d dapat dilihat seperti pada bagian berikut,

$$MNBC(d) = \operatorname{argmax}_{c \in C} P(c) \times \prod_{i < k < n_d} P(x_k | c) \quad (2)$$

dengan,

$P(c)$ = Probabilitas *prior* dari kelas (target)

$P(x_k | n)$ = Probabilitas kata ke- k dalam kelas c dibanding seluruh kelas lainnya.

C = Himpunan seluruh kelas yang dapat dimiliki d

n_d = Jumlah seluruh kata unik dalam d .

Akan tetapi, saat diimplementasikan, persamaan Multinomial Naïve Bayes Classifier (Persamaan 2) dapat menyebabkan kondisi *floating point underflow*, sehingga seringkali dilakukan tahap log normalization pada formula ini sehingga persamaannya berubah menjadi [15],

$$MNBC(d) = \operatorname{argmax}_{c \in C} \left[\log P(c) + \sum_{1 < k < n_d} \log P(x_k | c) \right] \quad (3)$$

Di sisi lain, algoritma Complement Naive Bayes Classifier (CNBC) merupakan algoritma yang akan melengkapi Algoritma MNBC. Pada algoritma CNBC proses pengklasifikasian kelas c untuk sebuah dokumen d didefinisikan oleh persamaan di bawah ini [6],

$$CNBC(d) = \operatorname{argmax}_{c \in C} \left[\log P(c) + \sum_{1 < k < n_d} \log P(x_k | c') \right] \quad (4)$$

dengan nilai $P(x_k | c')$ merepresentasikan probabilitas kata ke- k yang diketahui pada kelas bukan c .

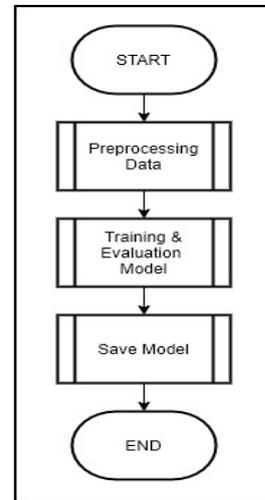
III. METODOLOGI DAN PENELITIAN

Dalam penelitian yang akan dilakukan, bagian Flowchart Umum akan digunakan untuk menjelaskan tahapan-tahapan yang akan dilakukan untuk mengetahui tingkat performa dari model gabungan algoritma MNBC dan CNBC. Kemudian bagian-bagian

lain di bawahnya akan digunakan untuk menjelaskan detail dari setiap tahap dalam Flowchart Umum.

A. Flowchart Umum

Dalam penelitian yang akan dilakukan detail dari proses-proses yang akan dilakukan ditunjukkan oleh gambar berikut.

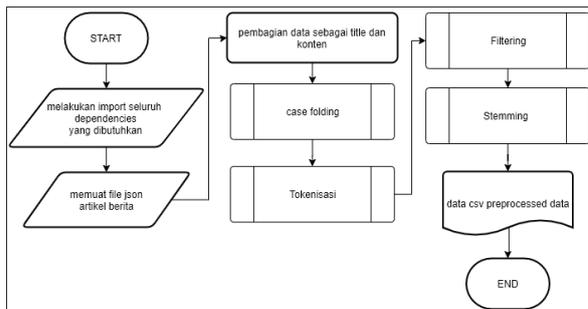


Gambar 1 Flowchart Umum

Sesuai dengan Gambar 1, tahap pertama yang dilakukan adalah *Preprocessing Data*. Kemudian, hasil dari proses akan digunakan untuk melakukan proses Model Training dan Evaluation. Pada proses ini akan dilakukan proses pelatihan model gabungan Algoritma Complement dan Multinomial Naïve Bayes Classifier menggunakan subset dari data yang tersedia. Setelah itu, berdasarkan model yang telah dilatih, proses evaluasi akan dilakukan dengan memberikan subset data lainnya yang belum pernah dikenali oleh model. Proses pelatihan dan evaluasi dari model akan dilakukan dengan mengujicobakan beberapa parameter dengan strategi *k-fold cross validation* dengan nilai k sama dengan 5. Terakhir, berdasarkan model dengan parameter terbaik, proses Save Model untuk menyimpan model data latih dan data uji yang terbaik.

B. Preprocessing Data

Dalam penelitian, data yang akan diolah terdiri dari 42.155 artikel dengan artikel terkait dengan Hiburan & Gaya Hidup sebesar 27.019, Olahraga sebesar 7.072, dan Indonesiaku sebesar 8.064. Pada proses Preprocessing Data, untuk setiap artikel, akan dilakukan tahap-tahap transformasi data tekstual seperti yang sudah dijelaskan pada bagian sebelumnya mengenai tahap-tahap Text Preprocessing. Urutan dari tahap-tahap Text Preprocessing yang akan dilakukan dalam penelitian digambarkan pada *flowchart* di Gambar 2.



Gambar 2 Flowchart Preprocessing Data

C. Training & Evaluation

Proses Training & Evaluation dimulai dengan memuat data hasil Text Preprocessing beserta label yang dimiliki oleh setiap data. Dalam proses Training & Evaluation, tahap yang pertama kali dilakukan adalah inisialisasi model klasifikasi dan TF-IDF. Dalam eksperimen yang dilakukan, terdapat 3 buah model yang akan dilatih dan dievaluasi yaitu model MNBC, model CNBC, serta model gabungan yang terdiri dari metode MNBC dan CNBC. Untuk setiap model akan dilakukan tahap pengujian parameter berikut.

1. Teks yang akan digunakan sebagai masukan dari model (analyser). Dalam penelitian yang ini, analyser yang dicobakan terdiri dari judul, konten serta gabungan dari judul dan konten.
2. Parameter model yang akan mempengaruhi bagaimana fitur-fitur akan dibentuk oleh metode TF-IDF. Dalam penelitian akan diujikan analyser berupa *word* dan *character*. Saat analyser di-*set* nilainya menjadi *word* maka metode TF-IDF akan membentuk representasi fitur dari dokumen berdasarkan kata-kata yang terdapat dalam sebuah dokumen. Di sisi lain, saat analyser di-*set* nilainya menjadi *character* maka fitur akan dibentuk berdasarkan huruf-huruf yang menyusun setiap kata dalam tiap dokumen.
3. Parameter min DF dan max DF yang akan memfilter jumlah kemunculan kata (*terms*) berdasarkan suatu *threshold* tertentu. Pada metode TF-IDF, fitur hanya akan dibentuk berdasarkan kata-kata yang jumlah kemunculannya berada di antara min DF dan max DF.
4. Parameter *ngram_range* yang mengatur rentang nilai dari kombinasi kata ataupun huruf yang akan dianggap sebagai fitur.

Berdasarkan setiap model yang telah dibentuk dari setiap kombinasi parameter yang telah diberikan, akan dilakukan proses evaluasi dengan menggunakan metrik F1-Score. Berdasarkan setiap parameter teks masukan yang diuji dan parameter model akan dicatat nilai parameter dari model dengan tingkat F1-Score terbaik. Visualisasi dari proses pelatihan dan evaluasi yang dilakukan dapat dilihat pada Gambar 3 pada halaman selanjutnya.

D. Save Model

Pada tahap ini, model dengan nilai F1-Score tertinggi yang dihasilkan pada tahap Train & Evaluation akan disimpan agar dapat digunakan oleh pihak terkait saat dibutuhkan.

IV. HASIL IMPLEMENTASI

A. F1-Score

Berdasarkan tahap-tahap metodologi penelitian yang telah dijelaskan pada bagian sebelumnya, Gambar 4 (pada halaman selanjutnya) menunjukkan model-model dengan F1-Score terbaik untuk setiap analyser dan model yang diujicobakan.

Berdasarkan Gambar 4, terlihat bahwa model terbaik hasil pelatihan adalah model gabungan CNBC & MNBC dengan parameter *character model*, analyser berupa content & title, Max_DF sebesar 0,9 dan Min_DF sebesar 0,025, dan nilai *ngram_range* sebesar (3,4) dengan F1-Score sebesar 90,13%.

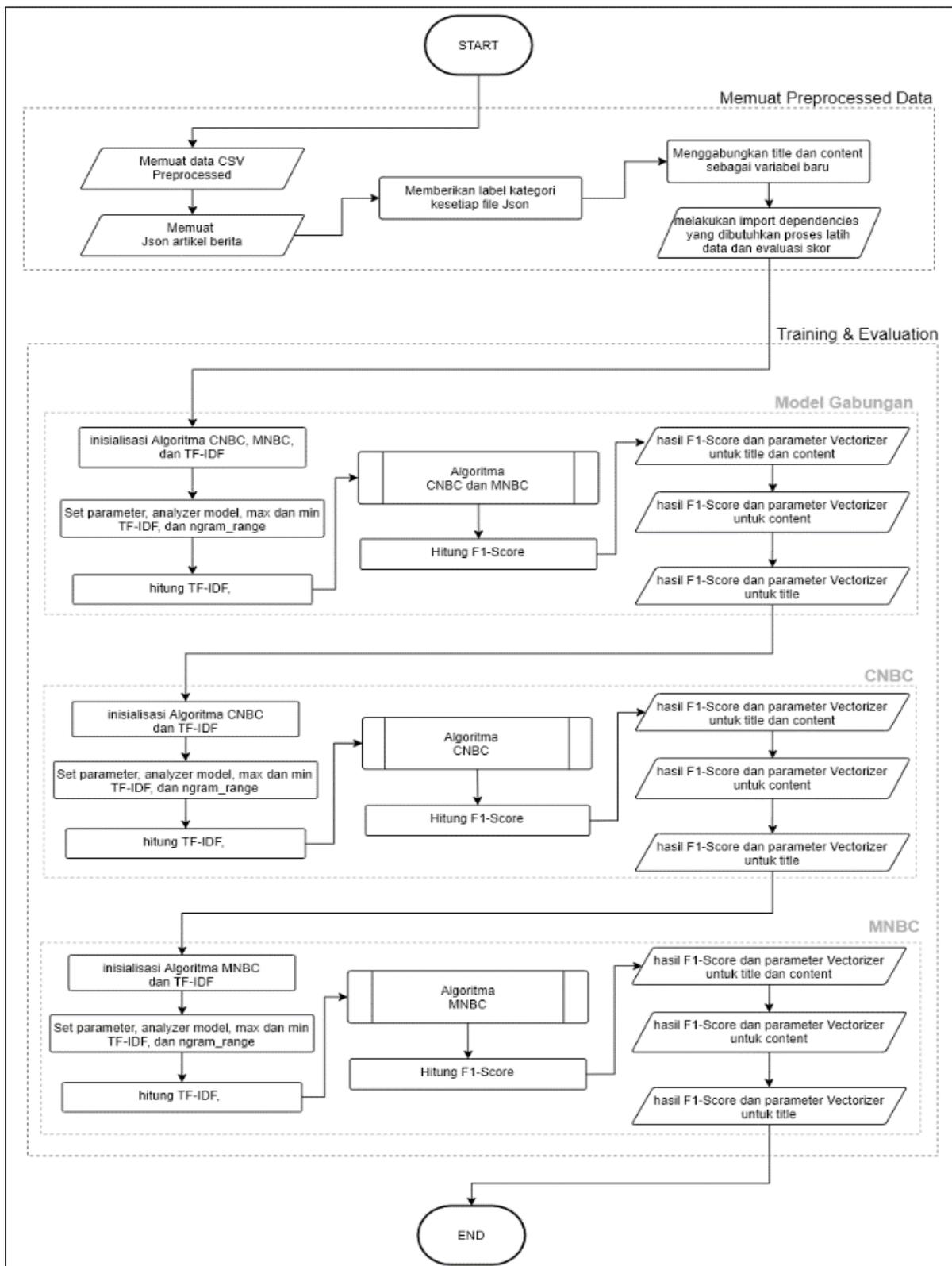
B. Confusion Matrix

Berdasarkan model dengan performa F1-Score terbaik yang dihasilkan, dilakukan analisis lanjutan dengan menggunakan metrik pengukuran lainnya yaitu, *confusion matrix* untuk mengetahui ketepatan prediksi dari model untuk setiap kategori berita. Proses analisis lanjutan dilakukan dengan melatih model baru dengan parameter yang serupa dengan model terbaik. Model ini dilatih dengan menggunakan 80% dari data artikel yang ada dan akan dievaluasi menggunakan 20% data artikel lainnya yang tersisa. Hasil confusion matrix dari model dari model terbaik dapat dilihat seperti pada Tabel 1.

Tabel 1 Confusion Matrix CNBC dan MNBC

		SISTEM		
		Hiburan & Gaya Hidup	Olahraga	Indonesiaku
AKTUAL	Hiburan & Gaya Hidup	4861	59	412
	Olahraga	55	1375	18
	Indonesiaku	330	8	1313

Melalui Tabel 1 dapat terlihat bahwa meskipun memiliki F1-Score yang sangat baik (90.13%), kesalahan prediksi paling kecil diperoleh oleh kategori Olahraga (sebesar 5% artikel dengan kategori Olahraga salah diprediksi) dan kesalahan prediksi paling besar diperoleh oleh kategori Indonesiaku (sebesar 20% artikel dengan kategori Indonesiaku salah diprediksi). Hal ini menunjukkan bahwa artikel dengan kategori Hiburan & Gaya Hidup dan Indonesiaku mungkin saja memiliki fitur yang serupa (cenderung tersusun berdasarkan kata-kata yang sama). Kemudian, berdasarkan confusion matrix yang dihasilkan, dihitung nilai precision, recall, serta F1-Score dari masing-masing kelas. Tabel 2 (halaman selanjutnya) menunjukkan nilai performa untuk masing-masing kategori.



Gambar 3 Flowchart Training & Evaluation Model

No	Methods	Model	Analyzer	Max_DF	Min_DF	N_GRAM_RANGE	F1-Score
1	CNBC & MNBC	Word Model	Content & Title	0,9	0,025	(1, 3)	89,28%
			Content	0,9	0,025	(1, 3)	89,23%
			Title	0,9	0,025	(1, 1)	63,37%
		Char Model	Content & Title	0,95	0,025	(3, 4)	90,13%
			Content	0,9	0,025	(3, 4)	90,05%
			Title	0,9	0,025	(2, 3)	71,27%
2	CNBC	Word Model	Content & Title	0,9	0,025	(1, 3)	88,55%
			Content	0,9	0,025	(1, 3)	88,55%
			Title	0,9	0,05	(1, 1)	62,86%
		Char Model	Content & Title	0,9	0,025	(3, 4)	89,53%
			Content	0,9	0,025	(3, 4)	89,45%
			Title	0,9	0,025	(2, 3)	66,31%
3	MNBC	Word Model	Content & Title	0,9	0,025	(1, 3)	89,15%
			Content	0,9	0,025	(1, 3)	89,08%
			Title	0,9	0,025	(1, 1)	62,99%
		Char Model	Content & Title	0,9	0,025	(3, 4)	90,03%
			Content	0,95	0,025	(3, 4)	89,96%
			Title	0,9	0,025	(2, 3)	68,71%

Gambar 4 Tabel Evaluasi F1-Score

Tabel 2 Rangkuman Perhitungan Rasio

Kategori	Precision	Recall	F1-Score
Hiburan & Gaya Hidup	0.93	0.91	0.92
Olahraga	0.95	0.95	0.95
Indonesiaku	0.75	0.80	0.77

Melalui Tabel 2 dapat terlihat bahwa model terbaik yang dihasilkan cenderung memiliki performa F1-Score yang baik saat memprediksi kelas Hiburan & Gaya Hidup dan Olahraga. Di sisi lain, untuk kelas Indonesiaku, model cenderung memiliki performa F1-Score yang kurang memuaskan. Hal ini mungkin saja disebabkan oleh jumlah data dari kelas Indonesiaku yang berjumlah sangat sedikit jika dibandingkan dengan jumlah data kelas lainnya (berbeda sampai dengan 992 artikel dengan kelas Olahraga dan berbeda hampir 20.000 data dengan kelas Hiburan & Gaya Hidup). Untuk meningkatkan performa dari kelas Indonesiaku tentunya dapat diujicobakan model-model lain yang bersifat lebih stabil untuk data yang bersifat *imbalance*.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil implementasi yang sudah dilakukan, maka penelitian dapat disimpulkan bahwa implementasi model algoritma Complement dan Multinomial Naïve Bayes Classifier dapat untuk mengklasifikasikan Kategori berita merahputih.com milik PT Merah Putih Media berhasil dilakukan. Berdasarkan percobaan yang telah dilakukan,

dihasilkan model terbaik dengan performa F1-Score sebesar 90.13%.

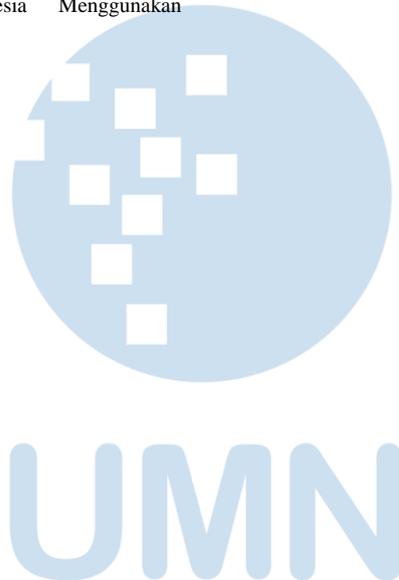
Berdasarkan penelitian yang telah dilakukan, terdapat beberapa saran untuk pengembangan penelitian lanjutan, antara lain.

1. Mencari model algoritma yang memiliki tingkat kompleksitas lebih tinggi. Dataset yang digunakan dalam penelitian ini memiliki jumlah yang besar tetapi model gabungan yang digunakan dalam penelitian memiliki tingkat kompleksitas yang cukup sederhana, sehingga apabila memiliki jumlah dataset yang besar dan memiliki model algoritma dengan tingkat kompleksitas tinggi dapat meningkatkan akurasi. Apabila ingin melakukan pengembangan penelitian, maka model yang disarankan adalah XGBoost Classifier atau Deep Neural Network Classifier.
2. Melakukan Entity Recognition. Model gabungan yang digunakan memiliki tingkat performa yang akurat. Alangkah baiknya jika dalam satu kategori, diketahui seluruh named-entity yang muncul, hal ini dapat memudahkan dan membantu proses Search Engine Optimization (SEO).

DAFTAR PUSTAKA

- [1] Y. Wibisono and L. M. Khodra, "Clustering Berita Berbahasa Indonesia," *FMIPA UPI & STEI ITB*, p. 32, 2005.
- [2] J. Irawan, Interviewee, *HRD PT Merah Putih Media*. [Interview]. 16 March 2020.

- [3] R. Feldman and J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*, New York: Cambridge University Press, 2007.
- [4] A. Rahman, Wiranto and A. Doewes, "Online News Classification using Multinomial Naive Bayes," *ITSMART: Jurnal Ilmiah Teknologi dan Informasi*, p. 32, 2017.
- [5] N. A. Siti, "KLASIFIKASI BERITA ONLINE MENGGUNAKAN METODE SUPPORT VECTOR MACHINE DAN K-NEAREST NEIGHBOR," INSTITUT TEKNOLOGI SEPULUH NOPEMBER, SURABAYA, 2016.
- [6] B. Herwijayanti, E. D. Ratnawati and L. Muflikhah, "Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, pp. 306-312, 2018.
- [7] D. M. J. Rennie, L. Shih, J. Teevan and R. D. Krager, "Tackling the Poor Assumptions of Naive Bayes Text Classifiers," *International Conference on Machine Learning*, pp. 1-8, 2003.
- [8] F. Sebastiani, "Machine Learning in Automated Text Categorization: a Bibliography," *ACM Computing Surveys*, pp. 1-47, 2001.
- [9] K. M. Dalal and A. M. Zaveri, "Automatic Text Classification: A Technical Review," *International Journal of Computer Applications*, pp. 37-40, 2011.
- [10] P. D. Langgeni, A. Z. Baizal and F. Y. A.W., "Clustering Artikel Berita Berbahasa Indonesia Menggunakan Unsupervised Feature Selection," *Seminar Nasional Informatika UPN*, pp. D1-D10, 2010.
- [11] S. M. Hudin, "IMPLEMENTASI METODE TEXT MINING DAN K-MEANS CLUSTERING UNTUK PENGELOMPOKAN DOKUMEN SKRIPSI (STUDI KASUS BRAWIJAYA)," UNIVERSITAS BRAWIJAYA, MALANG, 2018.
- [12] J. Asian, E. H. Williams and M. M. S. Tahaghoghi, "A Testbed for Indonesian Text Retrieval," *Australasian Document Computing Symposium*, pp. 1-4, 2004.
- [13] A. A. A. Putra, "Implementasi Text Summarization Menggunakan Metode Vector Space Model Pada Artikel Berita Berbahasa Indonesia," Perpustakaan UNIKOM, Bandung, 2016.
- [14] F. Azuaje, I. Witten and F. E., *Data Mining: Practical Machine Learning Tools and Techniques*, Northen Ireland: BioMedical Engineering OnLine , 2006.
- [15] D. C. Manning, P. Raghavan and H. Schutze, *Introduction to Information Retrieval*, Cambirdge: Cambridge University Press, 2008.
- [16] P. U. D. Singh, A. Tiwari and K. R. Singh, *Soft-Computing-Based Nonlinear Control Systems Design*, Madhya Pradesh: Madhav Institute of Technology and Science, 2018.



Sistem Konten Pembelajaran di Indonesia : *Systematic Literature Review*

Insanur Hanifuddin¹, Dhomas Hatta Fudholi², Sri Mulyati³

^{1,2,3} Program Studi Teknik Informatika, Universitas Islam Indonesia, Yogyakarta, Indonesia

¹17523022@students.uui.ac.id, ²hatta.fudholi@uui.ac.id, ³mulya@uui.ac.id

Diterima 13 Februari 2021

Disetujui 16 Juni 2021

Abstract—Wikipedia is the largest web-based digital encyclopedia today that contains almost all general knowledge. With Systematic Literature Review (SLR), this research examines Wikipedia, other encyclopedias and other online media that contain certain topics with target users in order to find out the potential and need to build a learning content system. The study was conducted on literature related to Wikipedia, encyclopedias, education and children's interests, especially at the elementary school level. Literature search was done by entering several main keywords in Google Scholar such as "Wikipedia", "encyclopedia", "elementary school curriculum", "educational content" and "learning media". Literature is also obtained through the official website of the Ministry of Education and Culture which contains education standards for elementary and junior high schools, educational assessment standards, as well as literacy and numeracy learning modules at the elementary level. The results of the literature analysis include 4 classifications based on topics, namely evaluation of use, content, online learning and media. Based on the results of the analysis of 30 papers, it was found 9 papers discussing content, 14 papers discussing content in specific fields, and 12 references discussing online media, so it is known that there has not been much research on digital encyclopedias for education.

Index Terms—education, encyclopedia, online learning, wikipedia

I. PENDAHULUAN

Wikipedia merupakan ensiklopedia terbesar di dunia dan merupakan salah satu situs web yang paling populer¹, dengan lebih dari lima ratus juta kunjungan per hari [1]. Dalam konteks pendidikan, ensiklopedia daring seperti Wikipedia memiliki potensi yang sangat baik untuk digunakan sebagai media pembelajaran yang sesuai kurikulum.

Media pembelajaran untuk anak-anak dapat diwujudkan dalam bentuk yang beragam, salah satunya adalah ensiklopedia. Ensiklopedia merupakan wadah dari kumpulan pengetahuan dan informasi yang pada umumnya ditampilkan dengan teks atau gambar. Media pembelajaran daring diharapkan dapat menjadi media untuk jangka panjang meskipun

pembelajaran secara tatap muka sudah dapat dilaksanakan kembali.

Media pembelajaran daring mendukung siswa dalam belajar secara mandiri dan mengkonstruksi pengetahuannya. Pada umumnya pembelajaran daring menggunakan media ruang guru [25], sehingga dari penelitian tersebut diketahui bahwa belum banyak yang memanfaatkan ensiklopedia daring seperti Wikipedia untuk menjadi platform pembelajaran secara spesifik.

Berdasarkan pencarian, ada satu literatur [2] yang membahas tentang perancangan ensiklopedia untuk anak usia 9-12 tahun tentang ekologi dan lingkungan tentang manusia dan lingkungannya, namun masih dalam bentuk buku. Belum banyak penelitian yang membahas tentang ensiklopedia digital, namun sudah ada satu penelitian mengenai ensiklopedia digital tentang dunia tumbuhan [3]. Dalam konteks pembelajaran daring ditemukan beberapa literatur yang relevan seperti literatur [4]–[6] namun tidak dalam bentuk ensiklopedia daring.

Penelitian yang dibangun akan menggunakan metode Systematic Literatur Review (SLR) , agar memudahkan dalam menemukan literatur sesuai dengan tujuan maka, literatur diklasifikasikan sesuai instrumen penelitian.

Padahal di jaman pandemik ini, setiap lapisan pendidikan melakukan pembelajaran secara daring. untuk itu Sistem Konten Pembelajaran disini disebut dengan Ensiklopedia daring penting untuk di buat untuk mendukung dalam proses pembelajaran, saat pandemi ini pelajar dituntut untuk mampu belajar mandiri sehingga sistem ini dapat digunakan sebagai bahan rekonstruksi pemahaman materi

Pada pendidikan tingkat dasar memiliki karakteristik pembelajaran yang menyeluruh dan belum terfokus pada bidang disiplin ilmu. yang bagi masyarakat umum sering dilupakan karena belum dianggap penting untuk itu kajian ini bertujuan untuk mengetahui hal hal penting dalam membangun ensiklopedia daring yang khusus untuk anak sekolah dasar dengan konten yang sesuai dengan kurikulum 2013.

¹ <https://ahrefs.com/blog/most-visited-websites/>

II. METODE

Pada bagian ini kajian pustaka dilakukan dengan metode Systematics Literature Review dengan mengadopsi pada [31]. Berikut ini gambaran dari tahapan penelitian :



Gambar 1. Tahapan Penelitian

Berikut ini penjelasan terkait tahapan penelitian yang dilakukan :

1. Pertanyaan Penelitian

Pertanyaan penelitian yang harapannya dapat menemukan tujuan penelitian adapun pertanyaan penelitiannya adalah:

- Bagaimana kondisi evaluasi pengguna terkait penggunaan media pembelajaran?
- Bagaimana cakupan dari bidang ilmu pada konten media pembelajaran
- Bagaimana peranan media pembelajaran yang diterapkan?
- Apa saja bentuk media pembelajaran yang digunakan?

2. Proses Pencarian

Pencarian literatur sesuai topik yang membahas tentang Wikipedia, ensiklopedia, pendidikan dasar dan media pembelajaran daring.

3. Pemilihan Bidang

Kemudian penulis melakukan pemilihan paper yang relevan berdasarkan paper yang diambil dengan mendefinisikan bahwa makalah harus menggambarkan media daring dan sistem konten pembelajaran. Pada jenjang pendidikan kami merujuk pada Permendikbud No 67 Tahun 2013 terkait kerangkadansar dan struktur kurikulum 2013 dan memilih untuk tingkat sekolah yang memiliki karakteristik pembelajaran menyeluruh dan belum terpusat pada salah satu bidang, maka jenjang pendidikan dasar yang sesuai.

4. Ekstraksi Data

Dari literatur tersebut dilakukan analisis sistem konten pembelajaran terhadap isi literatur dan diklasifikasikan menjadi empat, yakni evaluasi penggunaan, konten, pembelajaran daring dan media pembelajaran.

Hasil dari kajian ini diharapkan dapat menjadi rujukan untuk membangun ensiklopedia digital khusus anak dengan kurikulum 2013 sesuai standar

pendidikan di Indonesia, selanjutnya akan dijabarkan metode yang digunakan untuk pencarian literatur dan analisis literatur berdasarkan topik yang dijadikan acuan.

A. Metode Pencarian

Proses pencarian dilakukan pada Google Scholar dengan memasukkan kata kunci utama Wikipedia, ensiklopedia, kurikulum sekolah dasar, konten pendidikan dan media pembelajaran daring. Pencarian juga dilakukan pada situs resmi milik Kementerian Pendidikan dan Kebudayaan untuk menemukan kurikulum dan penilaian pendidikan tingkat dasar. Dari hasil pencarian dilakukan pemilihan literatur berdasarkan yang keseluruhannya mencakup:

- Literatur yang dipilih dipublikasikan dalam rentang tahun 2016 hingga 2020, yang mengindikasikan kebaruan referensi.
- Literatur membahas tentang poin penting yang dibutuhkan dalam pengembangan ensiklopedia dengan konten pendidikan.
- Literatur membahas tentang penggunaan dan potensi media daring untuk pembelajaran.
- Literatur berisi standar isi kurikulum dan standar penilaian pendidikan untuk sekolah tingkat dasar.
- Literatur dituliskan dalam Bahasa Inggris dan Bahasa Indonesia.

B. Metode Analisis

Analisis untuk literatur yang terpilih dilakukan dengan cara menjabarkan inti dari literatur tersebut dalam satu tabel untuk mengklasifikasikannya dalam empat topik. Keempat klasifikasi tersebut meliputi evaluasi penggunaan, konten, pembelajaran daring dan media. Dalam konteks evaluasi penggunaan, dipaparkan analisis mengenai penelitian yang terfokus dengan penggunaan, manfaat, potensi dan perspektif terhadap Wikipedia/ensiklopedia atau media lain. Dalam konteks konten, dipaparkan analisis terhadap topik khusus dari ensiklopedia maupun media lainnya yang masih terkait dengan pendidikan. Dalam konteks pembelajaran daring, dipaparkan analisis mengenai media daring apapun yang dimanfaatkan untuk kebutuhan pendidikan dan dampak dari penggunaan media daring sebagai sarana pendukung pendidikan. Sedangkan dalam konteks media pembelajaran, dipaparkan bentuk media yang digunakan dalam penelitian yang dikaji. Analisis yang dilakukan bertujuan untuk mempermudah mendapatkan informasi pendukung dari literatur-literatur terpilih untuk membangun ensiklopedia digital yang berisi pembelajaran untuk anak sekolah dasar sesuai kurikulum. Hasil dari analisis meliputi informasi yang esensial untuk mengembangkan ensiklopedia sebagai alat pembelajaran.

III. HASIL DAN PEMBAHASAN

Berdasarkan pencarian melalui Google Scholar dan situs resmi Kementerian Pendidikan dan Kebudayaan ditemukan 49 referensi, namun hanya 26 yang lebih sesuai dengan kebutuhan informasi untuk membangun ensiklopedia dengan konten pendidikan. Dari 49 referensi tersebut, terdapat 14 yang tidak memenuhi kriteria rentang tahun dan 9 referensi yang kurang sesuai dan tidak lengkap pembahasannya mengenai ensiklopedia sebagai media pembelajaran. Pada Tabel 1 dipaparkan perbandingan informasi yang diperoleh dari setiap referensi yang dikaji. Tabel 1 memberikan hasil analisis informasi terkait kesimpulan mengenai hal yang penting untuk membangun media pembelajaran.

Dari referensi tersebut, 9 dari 26 referensi mengandung informasi yang dibutuhkan mengenai Wikipedia/ensiklopedia yang meliputi penggunaan ensiklopedia sampai dengan persepsi dosen dan mahasiswa terhadap wikipedia. 13 dari 26 referensi

memaparkan informasi tentang kurikulum dan konten yang diangkat sebagai pembahasan dari referensi tersebut, dan 12 dari 26 referensi memaparkan informasi media pembelajaran daring yang meliputi penggunaan, efektifitas dan model media yang digunakan pada referensi tersebut. Pada poin 3.1 akan dilakukan analisis klasifikasi referensi dan digunakan sebagai penyaring informasi untuk membangun sistem konten pembelajaran : ensiklopedia daring sehingga menjadi bermanfaat, disukai dan dapat terus digunakan meskipun sekolah seperti biasa sudah dilaksanakan kembali seperti semula dan pada poin 3.2 akan dilakukan analisis terhadap *gap* dari referensi yang digunakan secara lebih detail.

A. Analisis Klasifikasi

Dari hasil kajian literatur, diperoleh empat klasifikasi utama yang digunakan sebagai penyaring informasi untuk membangun ensiklopedia daring dengan konten pendidikan yang efektif untuk anak sekolah dasar. Berikut ini hasil kajian literatur :

Tabel 1. Klasifikasi Literatur

Makalah	Tahun	Evaluai Penggunaan	Konten	Pembelajaran Daring	Media
[1]	2017	Alasan orang membaca Wikipedia.			Wikipedia.
[2]	2018	Media pembelajaran tentang ekologi dan lingkungan.	Ekologi manusia dan lingkungannya.		Buku.
[3]	2018	Ensiklopedia digital sebagai sumber belajar.	Tumbuhan tingkat tinggi.		Ensiklopedia Digital
[4]	2018		Pendidikan Agama Islam dan Budi Pekerti.	Media literasi dalam pembelajaran.	Beragam media pembelajaran.
[5]	2019		Pelajaran biologi.	Konten <i>e-learning</i> berbasis SCORM (<i>Shareable Content Object Reference Model</i>).	Microsoft Word dan Microsoft Power Point.
[6]	2018		Pendidikan Jasmani dan Olahraga.		Aplikasi <i>Mobile Learning</i> .
[7]	2017	Wikipedia sebagai situs warisan budaya.	Warisan budaya.	<i>Crowd Sourcing</i>	Wikipedia.
[8]	2017	Wikipedia sebagai sumber pengetahuan dari <i>Open-domain Question Answering System</i> .			
[9]	2018	Wikipedia sebagai sumber pengetahuan tentang penyakit.	Penyakit dan gejalanya.		
[10]	2016	Wikipedia sebagai media kolaborasi daring untuk pengetahuan bersama.		<i>Crowd Sourcing</i>	
[11]	2020	Analisis tingkat kesiapan pengguna ensiklopedia anak.		Guru sebagai kontributor.	Ensiklopedia Digital.
[12]	2018	Persepsi dosen dan mahasiswa terhadap penggunaan Wikipedia sebagai sumber informasi digital ilmiah.			
[13]	2020		.	Konten media sosial sebagai alternatif pembelajaran	Media sosial.
[14]	2017		Tematik.		Buku.
[15]	2016		Standar isi pendidikan dasar dan menengah.		
[16]	2016		Standar penilaian pendidikan.		
[17]	2017		Gerakan literasi nasional.		
[18]	2016		Gerakan literasi sekolah dasar.		
[19]	2020		Modul literasi dan	Program pembelajaran	<i>E-book</i> .

			numerasi.	jarak jauh.	
[20]	2020	Pemanfaatan teknologi informasi untuk pembelajaran daring.		Tantangan dan peluang pembelajaran daring pada masa pandemi Covid-19.	Edmodo, EdLink, Moodle, Google Classroom dan Schoology.
[21]	2020	Pemanfaatan kemajuan IT guna mendukung proses belajar mengajar.		Dampak positif pembelajaran daring di Indonesia.	
[22]	2020	Penggunaan aplikasi Ruang Guru terhadap minat belajar siswa SMA.		Minat belajar daring menggunakan aplikasi Ruang Guru.	Ruang Guru.
[23]	2020			Model dan desain pembelajaran daring.	Zoom, Google Classroom, Webex Meeting.
[24]	2020			Variasi pembelajaran daring.	Zoom, Google Classroom.
[25]	2018			Pengaruh komunikasi media daring aplikasi Ruangguru terhadap peningkatan prestasi belajar.	Ruang Guru.
[26]	2020			Kehadiran sosial Ruangguru dan kemampuan beradaptasi dengan kondisi yang terjadi.	Ruang Guru.
[27]	2017	Pengembangan media pembelajaran komik digital.	Ilmu Pengetahuan Sosial.		Komik Digital.
[28]	2018	Game android untuk meningkatkan minat belajar matematika.	Matematika.		Game.
[29]	2020	Pengaruh media pembelajaran daring dalam pemahaman dan minat belajar siswa.	Fisika.	Penerapan media pembelajaran daring.	
[30]	2017	Pengembangan aplikasi media pembelajaran game interaktif.	Bahasa Inggris.		Game.

Dari hasil analisis dalam sisi pemanfaatan atau penggunaan terdapat 16 penelitian yang memanfaatkan maupun mengembangkan berbagai media untuk penelitiannya. Terdapat 1 penelitian [1] yang memanfaatkan media Wikipedia untuk mengungkap mengapa orang mau membaca Wikipedia. Terdapat 9 penelitian yang memanfaatkan berbagai media untuk dikembangkan lebih lanjut seperti pada penelitian [2] yang mengembangkan media pembelajaran tentang ekologi manusia dan lingkungannya dalam bentuk buku, penelitian [3], [7]–[10] mengembangkan ensiklopedia dengan beragam konten di dalamnya. Terdapat 1 penelitian yang mengembangkan komik digital sebagai media pembelajaran ilmu pengetahuan sosial [27], dan 2 penelitian yang mengembangkan game sebagai media pembelajaran seperti pada penelitian [28] dan [30]. Terdapat 1 penelitian yang menganalisis tingkat kesiapan anak terhadap ensiklopedia digital [11] dan 5 penelitian yang membahas pemanfaatan dan penggunaan beragam media yang digunakan seperti pada penelitian [12] yang membahas persepsi dosen dan mahasiswa terhadap penggunaan Wikipedia, penelitian [20], [21] membahas pemanfaatan teknologi informasi dan penelitian [22], [29] yang menganalisis media terhadap minat belajar siswa.

Berdasarkan analisis terhadap pembelajaran daring ditemukan berbagai media yang digunakan pada pembelajaran daring seperti pada penelitian [4], berdasarkan hasil survei yang dilakukan pada

penelitian [20] menyatakan 50% mahasiswa belum memiliki laptop dan 80% mahasiswa menyatakan susah mendapatkan sinyal dan boros penggunaan paket data. Pada penelitian [24] ditemukan kelebihan dan kekurangan dari pembelajaran daring yang meliputi :

Kelebihan pembelajaran daring :

- Menyenangkan dan menarik.
- Efektif dan lebih mudah dimengerti.
- Belajar sambil bermain.
- Belajar hal baru secara daring.
- Tidak lelah dan bisa santai dalam belajar.
- Lebih semangat dalam belajar

Kekurangan pembelajaran daring :

- Siswa bosan belajar di rumah dan senang belajar di sekolah.
- Kesulitan dalam koneksi internet.
- Siswa tidak bisa berdiskusi dan bertemu secara langsung.
- Siswa lebih mengerti penjelasan dari guru secara langsung.
- Siswa merasa banyak beban tugas yang diberikan guru.
- Siswa merasa pusing dan lelah terlalu sering berhadapan dengan *gadget*.

B. Analisis Media

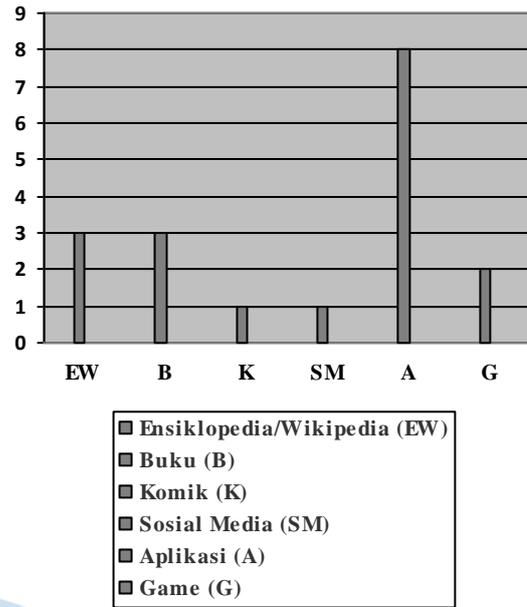
Hasil analisis diperoleh 6 media berbeda yang digunakan pada penelitian sebelumnya. Gambar 1 menunjukkan banyaknya media yang masih banyak digunakan pada penelitian sebelumnya.

Terdapat 3 penelitian yang memanfaatkan penggunaan media ensiklopedia seperti pada penelitian [3] yang mengembangkan ensiklopedia digital mengenai tumbuhan tingkat tinggi, penelitian [7] yang mengembangkan ensiklopedia mengenai warisan budaya dan penelitian [11] yang menganalisis tingkat kesiapan anak terhadap ensiklopedia digital.

Media dalam bentuk buku juga masih digunakan seperti pada penelitian [2] yang mengembangkan media pembelajaran tentang ekologi manusia dan lingkungannya. Ada satu penelitian yang mengembangkan media pembelajaran ilmu pengetahuan sosial dengan media komik digital yakni pada penelitian [27].

Sosial media juga menjadi salah satu media yang menjadi pilihan sebagai media pembelajaran, seperti pada penelitian [13] yang meneliti konten media sosial *Twitter*. Ada 2 penelitian yang memanfaatkan game sebagai media pembelajaran seperti pada penelitian [28] yang mengembangkan *game* pada *Android* untuk pembelajaran matematika dan pada penelitian [30] yang mengembangkan media pembelajaran *game* interaktif Bahasa Inggris.

Aplikasi menjadi media yang paling banyak dimanfaatkan maupun dikembangkan dari semua penelitian di atas. Terdapat 8 penelitian memanfaatkan atau mengembangkan aplikasi sebagai media dan fasilitas pendukung pembelajaran. Pada penelitian [5] memanfaatkan Microsoft Word dan Microsoft Power Point untuk pembelajaran daring berbasis SCORM (*Shareable Content Object Reference Model*). Pada penelitian [6] dikembangkan aplikasi *mobile learning* untuk membantu meningkatkan pembelajaran pendidikan jasmani dan olahraga. Terdapat 3 penelitian yang memanfaatkan aplikasi *Google Classroom* sebagai media pembelajaran seperti pada penelitian [20], [23] dan [24]. Terdapat 3 penelitian yang menggunakan Ruang Guru sebagai objek penelitiannya yang meneliti minat belajar daring [22], peningkatan prestasi belajar [25] dan kemampuan beradaptasi dari aplikasi Ruang Guru [26].



Gambar 2. Hasil analisis media

C. Analisis Gap Literatur

Pada analisis gap ini penulis akan membahas terkait perbandingan klasifikasi yang digunakan sebagai pertanyaan penelitian yang meliputi evaluasi pengguna, konten, pembelajaran daring, dan media berdasar referensi yang dihimpun terdapat potensi dari pemanfaatan ensiklopedia daring untuk mendukung pembelajaran. Hal ini ditunjukkan dari 30 penelitian sebelumnya didapatkan 9 referensi yang membahas mengenai Wikipedia dan ensiklopedia dalam konteks penggunaan maupun pengembangan. Berdasarkan paparan Tabel 1 di atas, didapatkan alasan orang mengakses Wikipedia karena akses yang mudah dan ketersediaan informasi yang lengkap. Banyak penelitian sebelumnya yang berfokus pada kolaborasi dalam pemanfaatan Wikipedia untuk dikembangkan dengan konten tertentu. Ada satu penelitian yang mengembangkan ensiklopedia untuk anak, namun masih dalam bentuk buku dan belum dalam bentuk media digital.

Dalam konteks konten didapatkan 14 referensi yang mencakup pembahasan mengenai bidang spesifik dan 2 referensi mengenai peraturan pemerintah yang mengatur tentang standar isi kurikulum dan standar penilaian untuk pendidikan sekolah tingkat dasar. Belum ada penelitian yang membahas media pembelajaran dengan konten yang sesuai dengan kurikulum yang berlaku di Indonesia.

Sedangkan dalam konteks pembelajaran daring, didapatkan 12 referensi yang mengangkat topik mengenai media daring. Media daring tidak hanya digunakan sebagai media pembelajaran, dapat juga dimanfaatkan untuk hal lain. Penelitian yang membahas tentang media daring banyak memaparkan kelebihan dan kekurangan dari pembelajaran daring

itu sendiri, selain itu ada juga yang membahas tentang tingkat kesiapan pengguna dan dampak yang ditimbulkan dari penggunaan media daring.

Berdasarkan hasil analisis, belum ada penelitian yang membahas semua klasifikasi pada Tabel 1 sekaligus, belum banyak penelitian tentang ensiklopedia digital untuk pendidikan dan dapat disimpulkan bahwa pembelajaran daring memiliki potensi tinggi untuk meningkatkan minat belajar anak asalkan konten yang ditujukan untuk anak-anak tepat sasaran. Dengan mempertimbangkan hal hal penting yang mendukung pada pembuatan konten pembelajaran dapat menciptakan sistem konten sistem pembelajaran yang sesuai untuk mendukung pembelajaran, pembelajaran daring menggunakan media ensiklopedia daring diharapkan dapat meningkatkan minat belajar bagi anak dan pembelajaran daring menjadi kebiasaan baru untuk jangka waktu yang panjang.

IV. SIMPULAN

Ensiklopedia sebagai media pembelajaran daring memiliki potensi yang baik untuk menjadi alat bantu bagi siswa sekolah, terutama pada tingkat dasar. Hal ini didapatkan dari analisis yang dapat kita lihat hubungan keterkaitan dari klasifikasinya sebagai berikut :

- Melihat Potensi dari evaluasi pengguna
- Karakteristik bidang ilmu dari konten
- Peran media belajar dari Pembelajaran daring
- Potensi Media dari analisis media

Ensiklopedia digital dapat menjadi pengganti buku sekolah bagi anak mengingat perkembangan teknologi saat ini yang sangat pesat. Untuk membangun ensiklopedia digital yang baik untuk siswa pendidikan tingkat dasar dan dapat bertahan lama dibutuhkan kolaborasi antara guru-guru di sekolah untuk membuat konten mengenai mata pelajaran yang diampu pada ensiklopedia tersebut. Konten mata pelajaran harus dibuat semenarik mungkin agar dapat meningkatkan minat belajar dan prestasi akademik anak. Konten juga harus dibuat kreatif dan tidak membosankan agar anak terus tertarik untuk belajar hal-hal baru.

Melalui ensiklopedia digital, siswa diharapkan dapat menerima model baru pembelajaran daring dengan konten yang disediakan melalui teks, gambar dan video. Peran orang tua juga menjadi hal yang harus diperhatikan dalam membimbing anak menggunakan perangkat untuk mengakses ensiklopedia digital tersebut agar anak tidak salah menggunakan perangkat untuk mengakses hal-hal yang negatif. Ensiklopedia digital dengan konten pendidikan belum banyak dikembangkan, terutama di Indonesia, sudah ada beberapa penelitian yang memanfaatkan media untuk pembelajaran namun

belum sesuai dengan kurikulum sekolah. Ensiklopedia digital tidak akan bisa dimanfaatkan dengan maksimal tanpa peran guru dan orang tua.

DAFTAR PUSTAKA

- [1] P. Singer *et al.*, "Why we read Wikipedia," *26th Int. World Wide Web Conf. WWW 2017*, pp. 1591–1600, 2017, doi: 10.1145/3038912.3052716.
- [2] G. D. Prameswari Sutanto, B. R. Noviadji, and C. B. Susilo, "Perancangan Buku Ensiklopedia Ekologi & Lingkungan Tentang Manusia dan Lingkungannya Untuk Anak Usia 9-12 Tahun," *Artika*, vol. 2, no. 1, pp. 1–11, 2018, doi: 10.34148/artika.v2i1.71.
- [3] K. Azis, M. Wiharto, and S. Saenab, "Ensiklopedia Digital Tumbuhan Tingkat Tinggi Dengan Fitur QR Code Sebagai Sumber Belajar Dunia Tumbuhan," *Biol. Teach. Learn.*, vol. 1, no. 2, pp. 109–120, 2018.
- [4] U. Wahidin, "Implementasi Literasi Media dalam Proses Pembelajaran Pendidikan Agama Islam dan Budi Pekerti," *Implementasi Literasi Media Dalam Proses Pembelajaran Pendidik. Agama Islam Dan Budi Pekerti*, pp. 229–244, 2018.
- [5] S. H. Bariah and S. M. Sidik, "Penerapan Konten E-learning Berbasis SCORM Untuk Meningkatkan Efektivitas Pembelajaran," vol. 5, pp. 1–10, 2019.
- [6] Y. Satria, S. Saptadi, H. Prastawa, and P. Jasmani, "Pengembangan Aplikasi Mobile Learning Sebagai Pelengkap Pembelajaran SMA Negeri 2 Semarang (Studi Kasus: Mata Pelajaran Pendidikan Jasmani dan Olahraga)," pp. 1–12, 2018.
- [7] C. Pentzold, E. Weltevrede, M. Mauri, D. Laniado, A. Kaltenbrunner, and E. Borra, "Digging Wikipedia: The online encyclopedia as a digital cultural heritage gateway and site," *J. Comput. Cult. Herit.*, vol. 10, no. 1, pp. 1–19, 2017, doi: 10.1145/3012285.
- [8] D. Chen, A. Fisch, J. Weston, and A. Bordes, "Reading Wikipedia to answer open-domain questions," *ACL 2017 - 55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap.)*, vol. 1, pp. 1870–1879, 2017, doi: 10.18653/v1/P17-1171.
- [9] E. P. G. Del Valle, G. L. Garcia, L. Prieto Santamaria, M. Zanin, E. M. Ruiz, and A. R. González, "Evaluating Wikipedia as a Source of Information for Disease Understanding," *Proc. - IEEE Symp. Comput. Med. Syst.*, vol. 2018-June, pp. 399–404, 2018, doi: 10.1109/CBMS.2018.00076.
- [10] J. Yun, S. H. Lee, and H. Jeong, "Intellectual interchanges in the history of the massive online open-editing encyclopedia, Wikipedia," *Phys. Rev. E*, vol. 93, no. 1, 2016, doi: 10.1103/PhysRevE.93.012307.
- [11] R. D. Kristy, E. D. Wahyuni, and N. Hayatin, "Analisis Tingkat Kesiapan Pengguna Ensiklopedia Anak Dengan Menggunakan Metode Technology Readiness Index," *Repositor*, vol. 2, no. 2, pp. 129–136, 2020.
- [12] E. Syafrita, "Persepsi Dosen dan Mahasiswa Program Studi Informasi Perpustakaan Dan Kearsipan Terhadap Wikipedia Sebagai Sumber Informasi Digital," no. September, pp. 99–103, 2018.
- [13] A. R. Randisa and A. Nurmandi, "Analisis Konten Media Sosial Twitter Sarana Pendidikan di Indonesia Study Kasus Ruang Guru," *J. Ilm. Tata Sejuta STIA Mataram*, vol. 6, no. 2, pp. 291–601, 2020, doi: 10.32666/tatasejuta.v6i2.135.
- [14] Suyono, T. Harsiati, and I. S. Wulandari, "Implementasi Gerakan Literasi Sekolah Pada Pembelajaran Tematik di Sekolah Dasar," *Sekol. Dasar Kaji. Teor. dan Prakt. Pendidik.*, pp. 116–123, 2017.
- [15] Menteri Pendidikan dan Kebudayaan, *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 21 Tahun 2016 Tentang Standar Isi Pendidikan Dasar*

- dan Menengah. Indonesia, 2016.
- [16] Menteri Pendidikan dan Kebudayaan, *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 23 Tahun 2016 Tentang Standar Penilaian Pendidikan*. Indonesia, 2016.
- [17] Kementerian Pendidikan dan Kebudayaan, *Panduan Gerakan Literasi Nasional*. Jakarta: Kementerian Pendidikan dan Kebudayaan, 2017.
- [18] Kementerian Pendidikan dan Kebudayaan and Direktorat Jenderal Pendidikan Dasar dan Menengah, *Panduan Gerakan Literasi Sekolah di Sekolah Dasar*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah, Kementerian Pendidikan dan Kebudayaan, 2016.
- [19] Pusmenjar, *Modul Belajar Siswa Kelas 1 Tema 1 Diriku Sub Tema 1 Keluarga, Modul Belajar Literasi dan Numerasi Jenjang SD Program Pembelajaran Jarak Jauh*, Modul. Jakarta: Kementerian Pendidikan dan Kebudayaan, 2020.
- [20] N. K. Suni Astini, "Tantangan Dan Peluang Pemanfaatan Teknologi Informasi Dalam Pembelajaran Online Masa Covid-19," *Cetta J. Ilmu Pendidik.*, vol. 3, no. 2, pp. 241–255, 2020, doi: 10.37329/cetta.v3i2.452.
- [21] Y. Pujilestari, "Dampak Positif Pembelajaran Online Dalam Sistem Pendidikan Indonesia Pasca Pandemi Covid-19," *Adalah*, vol. 4, no. 1, pp. 49–56, 2020, [Online]. Available: <http://journal.uinjkt.ac.id/index.php/adalah/article/view/15394/7199>.
- [22] Y. Violita, "Penggunaan Aplikasi Ruang Guru terhadap Minat Belajar Siswa SMA (Studi Korelasional Pengaruh Penggunaan Aplikasi Ruangguru terhadap Minat Belajar Siswa SMA Negeri 1 Berastagi)," 2020.
- [23] P. Marbun, "Disain Pembelajaran Online Pada Era Dan Pasca Covid-19," *CSRID J.*, vol. 12, no. 2, pp. 129–142, 2020, [Online]. Available: <http://csrid.potensi-utama.ac.id/index.php/CSRID/article/view/408>.
- [24] E. Pujiasih, "Membangun Generasi emas dengan Variasi Pembelajaran Online di Masa Pandemi Covid-19," *Ideguru J. Karya Ilm. Guru*, vol. 5, no. 1, pp. 42–48, 2020.
- [25] E. Langi, L. N. S., and L. Tulung, "Pengaruh Komunikasi Media Online Aplikasi Ruang Guru Pt. Ruang Raya Indonesia Terhadap Peningkatan Prestasi Belajar Siswa Sma Negeri 1 Manado," *Acta Diurna Komun.*, vol. 1, no. 3, pp. 1–16, 2018, [Online]. Available: <https://ejournal.unsrat.ac.id/index.php/actadiurnakomunikasi/article/view/25077/24778>.
- [26] R. A. Fattah and F. K. Sujono, "Social Presence of Ruangguru in Social Media during Covid-19 Pandemic," *J. Messenger*, vol. 12, no. 2, p. 180, 2020, doi: 10.26623/themessenger.v12i2.2276.
- [27] E. Sukmanasa, T. Windiyani, and L. Novita, "Pengembangan Media Pembelajaran Komik Digital Pada Mata Pelajaran Ilmu Pengetahuan Sosial Bagi Siswa Kelas V Sekolah Dasar Di Kota Bogor," *J. Pendidik. Sekol. Dasar*, vol. 3, no. 2, p. 171, 2017, doi: 10.30870/jpsd.v3i2.2138.
- [28] M. S. Itqan, "Pendekatan Game Android Untuk Meningkatkan Minat Belajar Matematika Siswa Sekolah Dasar," *J. EduMatSains*, vol. 2, no. 2, pp. 161–170, 2018, [Online]. Available: <http://ejournal.uki.ac.id/index.php/edumatsains/article/view/604>.
- [29] S. Maulidina and Y. B. Bhakti, "Pengaruh Media Pembelajaran Online Dalam Pemahaman Dan Minat Belajar Siswa Pada Konsep Pelajaran Fisika," *ORBITA J. Kajian, Inov. dan Apl. Pendidik. Fis.*, vol. 6, no. 2, p. 248, 2020, doi: 10.31764/orbita.v6i2.2592.
- [30] A. R. Nursifa, A. Orinda, and S. S. Pritanti, "Media Pembelajaran Interaktif Bahasa Inggris Kelas V Sekolah Dasar Negeri 02 Cinere," *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 1, p. 9, 2017, doi: 10.24014/coreit.v3i1.3381.
- [31] H. Hendrik, dan A. Hamzah. Flipped Classroom in Programming Course: A Systematic Literature Review. *International Journal Of Emerging Technologies In Learning (IJET)*, 16(02), pp. 220-236. 2021. doi:<http://dx.doi.org/10.3991/ijet.v16i02.15229>

UMN

Analisis *Clustering* Pengelompokan Penjualan Paket Data Menggunakan Metode K-Means

Dimas Galang Ramadhan¹, Indri Prihatini², Febri Liantoni³

^{1,2,3} Pendidikan Teknik Informatika dan Komputer, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sebelas Maret

Diterima 01 April 2021

Disetujui 13 Juni 2021

Abstract— At present, the COVID-19 pandemic requires all activities based in the network, starting from work, college, school, and everything based on the network. Certain *provider* users will experience excessive data plan usage. This also affects a counter that sells data packages, which must provide several data package services following current conditions. This research was conducted to analyze the grouping of sales of data packages that are often purchased by customers in a counter by using the K-Means method. The K-Means method is used because the K-Means algorithm is not influenced by the order of the objects used, this is proven when the writer tries to determine the initial *cluster* center randomly from one of the objects in the first calculation. sales of data packages at a counter. Variables used include Price, Active period, and several data packages. The K-Means *Cluster* Analysis algorithm is basically applied to the problem of understanding consumer needs, identifying the types of data package products that are often purchased. The K-Means algorithm can be used to describe the characteristics of each group by summarizing a large number of objects so that it is easier.

Index terms—K-Means, *Clustering*, Data Mining

I. PENDAHULUAN

Salah satu langkah dalam knowledge discovery databases adalah Data Mining. Data Mining merupakan teknik analisis data untuk mendapatkan informasi yang tersembunyi dari data yang kompleks dalam jumlah besar. Salah satu teknik analisa Data Mining adalah *Cluster* Analysis (Analisis Kelompok) atau dikenal sebagai *Clustering*. *Clustering* merupakan sebuah metode analisis dengan tujuan mengelompokkan data berdasarkan karakteristik objek pada suatu wilayah yang sama. Algoritma K-Means digunakan dalam mengembangkan metode *Clustering*, algoritma ini merupakan algoritma pengelompokan data secara nonhierarki (sekatan) yang membagi data kedalam dua atau lebih kelompok (*cluster*) dengan karakteristik sama.

Pada penelitian sebelumnya metode *clustering* digunakan untuk mengelompokkan daun berdasarkan fitur moment invarian. Pada penelitian ini metode *clustering* yang digunakan yaitu *Hierarchical Clustering* dan K-Means [1], [2].

Konter merupakan salah satu bidang usaha penjualan pulsa dan paket data. Dimana untuk situasi pandemi sekarang ini paket data banyak sekali dicari untuk kebutuhan aktivitas dalam jaringan, penggunaan *provider* dan jumlah paket data yang sering digunakan konsumen membuat bidang usaha ini harus menentukan jenis *provider*, jumlah paket data dan harganya yang akan diperbanyak stoknya berdasarkan penggunaan pelanggan. Oleh karena itu diperlukan data yang cukup agar bisa dianalisis lebih lanjut dan menghasilkan sebuah informasi yang tepat.

Berdasarkan hal tersebut diharapkan sebuah konter dapat menentukan suatu pengelompokan untuk menentukan *provider*, jumlah paket data, serta harga yang akan di perbanyak stoknya. Penulis juga mengupayakan agar bisa memberikan rangsangan kepada pengolah data supaya bisa mengumpulkan informasi lebih banyak yang dibutuhkan oleh sebuah konter, sehingga bisa menghasilkan pilihan yang tepat untuk menjawab kebutuhan dari sebuah Konter. Hasil penelitian ini akan memberikan informasi pilihan paket favorit yang diminati.

II. METODE PENELITIAN

A. *Clustering*

Clustering adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa group berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya [2], [3]. *Cluster* adalah sekelompok atau sekumpulan objek-objek data yang similar satu sama lain dalam *cluster* yang sama dan disimilari terhadap objek-objek yang berbeda *cluster* [4][1].

Data dari *clustering* memiliki tujuan, yaitu untuk meminimalisasikan objektif function yang diatur dalam proses *clustering*, dimana pada umumnya berusaha meminimalisasikan variasi dalam suatu *cluster* dan antar *cluster*.

B. Algoritma Klasterisasi K-Means

Metode algoritma K-Means digunakan pada penelitian ini untuk menentukan kelompok (*cluster*) yang paling baik untuk pemilihan *provider*, jumlah paket data, serta harga yang akan diperbanyak di Konter Plus Ultra Pabelan. Sehingga peluang untuk menambah

penghasilan penjualan semakin banyak dengan menambah stok *provider* yang paling banyak dibeli.

Metode K-Means merupakan metode pengelompokan data berdasarkan titik pusat *cluster* (centroid) [5]. Metode K-Means bertujuan untuk mengelompokan data dengan karakteristik yang sama dalam satu *cluster*.

Langkah awal proses *clusterisasi* data dengan menggunakan metode K-Means adalah penentuan titik pusat awal c_j , pada umumnya penentuan titik awal centroid didapatkan secara acak [5], [6]. Banyaknya centroid c_j yang didapatkan sesuai dengan jumlah *cluster* yang ditentukan di awal. Setelah titik pusat awal terbentuk kemudian langkah selanjutnya menghitung jarak tiap data x_i dengan titik pusat awal ke- j sampai k , dinotasikan dengan $d(x_i, c_j)$.

Kelanjutan dari jarak data dengan titik pusat tersebut dicari jarak terdekat sehingga data akan dikelompokkan berdasarkan centroid terdekat [7]. Langkah selanjutnya yaitu memperbarui titik centroid dengan cara menghitung rata-rata jarak dari seluruh data dengan centroid. Kemudian kembali lagi ke langkah awal. Iterasi ini dilakukan berulang-ulang hingga mendapatkan titik centroid sudah tidak berubah (konstan).

Metode K-Means digunakan tergantung dari data yang ada dan konklusi yang ingin dicapai dengan aturan sebagai berikut : (a) Jumlah *Cluster* perlu diinputkan. (b) Hanya memiliki atribut numeric [8].

Metode K-Means awalnya mengambil beberapa komponen dari populasi sebagai pusat titik awal atau pusat *cluster* awal [9]. Pada tahapan ini pusat *cluster* ditentukan secara acak serta menandai komponen tersebut ke dalam salah satu pusat *cluster* yang telah didefinisikan, berdasarkan hasil dari jarak minimum data dengan titik pusat *cluster*. Titik pusat awal (*cluster*) selanjutnya dihitung kembali hingga semua komponen data dikelompokkan kedalam masing-masing pusat *cluster* sehingga terbentuk posisi pusat *cluster* yang baru.

Terdapat 2 proses pada algoritma K-Means yaitu proses pencarian anggota dari tiap-tiap *cluster* dan proses pendeteksian lokasi pusat tiap *cluster*. Cara kerja algoritma K-Means (a) Menentukan k sebagai jumlah *cluster* yang ingin dibentuk. (b) Bagkitkan k titik pusat *cluster* awal secara acak. (c) Hitung jarak masing –masing data terhadap titik pusat *cluster*. (d) Pilih setiap data dengan titik pusat *cluster* terdekat. (e) Tentkan titik pusat *cluster* baru dengan menghitung rata-rata data yang terletak pada titik awal *cluster* yang sama. (f) Kembali pada langkah ke-3 jika posisi titik awal *cluster* lama dengan yang baru masih berbeda.

III. HASIL DAN PEMBAHASAN

Pada penelitian ini, terdapat 40 data yang digunakan dalam uji coba, dari 40 hasil data tersebut disusun pada tabel yang terdiri dari nama pembeli, besar kuota, masa aktif, dan harga.

Tabel 1. Data uji coba pembelian paket data

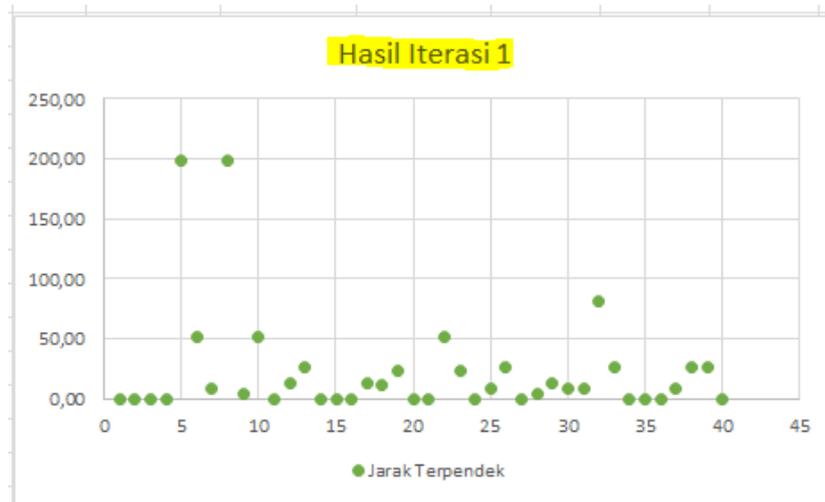
No	Nama Pembeli	Besar Kuota	Masa Aktif	Harga	No	Nama Pembeli	Besar Kuota	Masa Aktif	Harga
1	Irene	2,75	7	10	21	Carol	6	7	35
2	Liam	10	30	50	22	Anna	25	30	100
3	Austin	10	30	50	23	Blake	2	30	15
4	Tracey	2,75	7	10	24	Christopher	2,75	7	10
5	Mary	70	30	239	25	Brian	10	30	59
6	Tracey	25	30	100	26	Rose	8	30	76
7	Nicholas	10	30	59	27	Felicity	6	7	35
8	Christopher	70	30	239	28	Justin	3	7	14
9	Deirdre	3	7	14	29	Richard	2	30	39
10	Austin	25	30	100	30	Luke	10	30	59
11	Victor	10	30	50	31	Theresa	10	30	59
12	Samantha	2	30	39	32	Ava	30	30	129
13	Ryan	8	30	76	33	Harry	8	30	76
14	James	10	30	50	34	Amanda	10	30	50
15	Brian	6	7	35	35	Evan	10	30	50
16	Thomas	10	30	50	36	Stewart	2,75	7	10
17	Dan	2	30	39	37	Hannah	10	30	59
18	Joseph	3	30	41	38	Victor	8	30	76
19	Ryan	2	30	15	39	Nicola	18	30	75

20	Mary	6	7	35	40	Olivia	6	7	35
----	------	---	---	----	----	--------	---	---	----

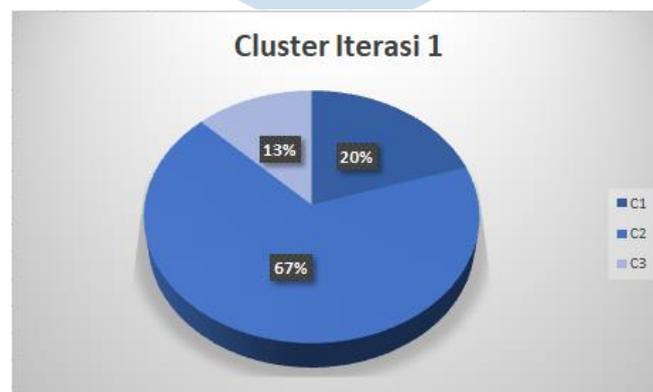
Dari 40 data tersebut kemudian ditentukan 3 *cluster* yang digunakan sebagai acuan dalam iterasi awal pada metode k-means yang digunakan. *Cluster* ini mewakili besaran paket yang umum tersedia, yaitu paket banyak, sedang dan sedikit. 3 *cluster* tersebut seperti ditunjukkan pada Tabel 2. Berdasarkan proses *clustering* yang dilakukan pada iterasi 1, maka diperoleh data iterasi 1 yang ditunjukkan dengan grafik iterasi 1 seperti pada Gambar 1 dan Gambar 2.

Tabel 2. Penentuan *Cluster* untuk Iterasi 1

No	Cluster	Besar Kuota	Masa Aktif	Harga	Jenis
1	C1	2,75	7	10	Sedikit
2	C2	10	30	50	Banyak
3	C3	6	7	35	Sedang



Gambar 1. Grafik data hasil iterasi 1

Gambar 2. Hasil *Cluster* Iterasi 1

Algoritme pengelompokan K-means dapat ditingkatkan secara signifikan dengan menggunakan teknik inisialisasi yang lebih baik, dan dengan mengulang (memulai kembali) algoritme tersebut [10].

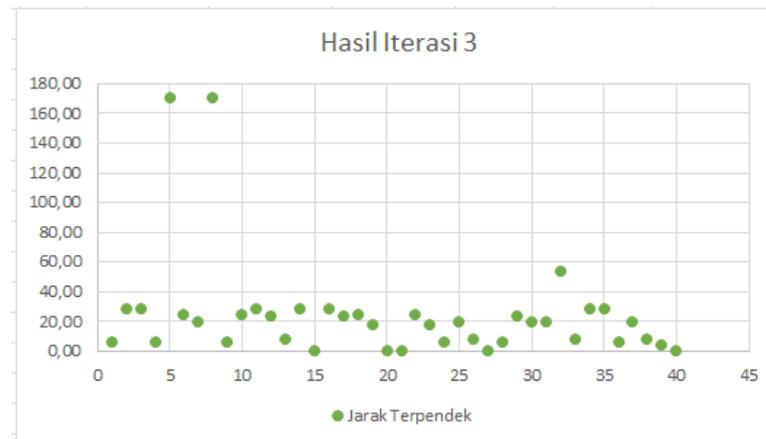
Jika data memiliki *cluster* yang tumpang tindih, k-means dapat meningkatkan hasil teknik inisialisasi. Ketika data memiliki *cluster* yang terpisah dengan baik, kinerja k-means sepenuhnya bergantung pada kebaikan inisialisasi.

Iterasi pada metode K-Means akan selalu dilakukan, sampai mendapatkan hasil yang optimal.

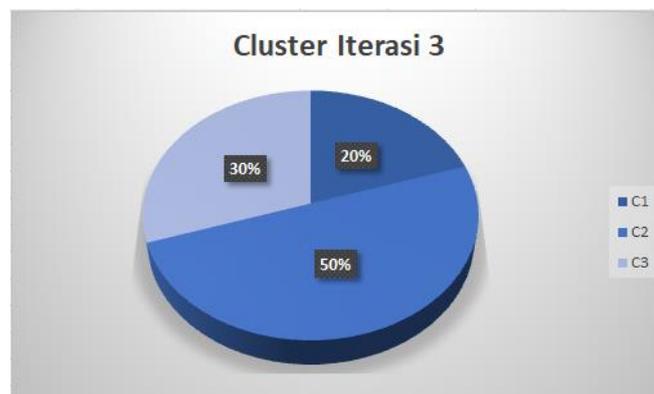
Pada proses metode K-Means pada literasi ke 2, diperoleh hasil yang belum optimal sehingga perlu dilakukan proses *clustering* perhitungan iterasi ke 3, dari hasil iterasi ke 3 diperoleh hasil yang optimal dari data uji coba yang digunakan. Dari hasil iterasi ke 3 ini akan dijadikan rujukan dalam penentuan hasil *cluster* yang dibentuk. Hasil penghitungan keseluruhan dari iterasi ke 3 ditunjukkan pada Tabel 3. Sedangkan grafik hasil iterasi ke 3 dan hasil *cluster* yang terbentuk seperti ditunjukkan pada Gambar 3 dan Gambar 4.

Tabel 3. Hasil penghitungan iterasi ke 3

No	Nama Pembeli	Besar Kuota	Masa Aktif	Harga	C1	C2	C3	Jarak Terpendek
1	Irene	2,75	7	10	6,18	72,36	25,21	6,18
2	Liam	10	30	50	42,15	27,96	27,75	27,75
3	Austin	10	30	50	42,15	27,96	27,75	27,75
4	Tracey	2,75	7	10	6,18	72,36	25,21	6,18
5	Mary	70	30	239	237,17	170,51	215,04	170,51
6	Tracey	25	30	100	92,18	24,47	71,52	24,47
7	Nicholas	10	30	59	50,37	19,23	33,48	19,23
8	Christopher	70	30	239	237,17	170,51	215,04	170,51
9	Deirdre	3	7	14	6,02	68,60	21,21	6,02
10	Austin	25	30	100	92,18	24,47	71,52	24,47
11	Victor	10	30	50	42,15	27,96	27,75	27,75
12	Samantha	2	30	39	31,84	40,74	23,69	23,69
13	Ryan	8	30	76	66,26	7,82	47,05	7,82
14	James	10	30	50	42,15	27,96	27,75	27,75
15	Brian	6	7	35	23,71	49,18	0,00	0,00
16	Thomas	10	30	50	42,15	27,96	27,75	27,75
17	Dan	2	30	39	31,84	40,74	23,69	23,69
18	Joseph	3	30	41	33,53	38,52	23,96	23,96
19	Ryan	2	30	15	17,48	63,86	30,74	17,48
20	Mary	6	7	35	23,71	49,18	0,00	0,00
21	Carol	6	7	35	23,71	49,18	0,00	0,00
22	Anna	25	30	100	92,18	24,47	71,52	24,47
23	Blake	2	30	15	17,48	63,86	30,74	17,48
24	Christopher	2,75	7	10	6,18	72,36	25,21	6,18
25	Brian	10	30	59	50,37	19,23	33,48	19,23
26	Rose	8	30	76	66,26	7,82	47,05	7,82
27	Felicity	6	7	35	23,71	49,18	0,00	0,00
28	Justin	3	7	14	6,02	68,60	21,21	6,02
29	Richard	2	30	39	31,84	40,74	23,69	23,69
30	Luke	10	30	59	50,37	19,23	33,48	19,23
31	Theresa	10	30	59	50,37	19,23	33,48	19,23
32	Ava	30	30	129	121,15	53,57	99,70	53,57
33	Harry	8	30	76	66,26	7,82	47,05	7,82
34	Amanda	10	30	50	42,15	27,96	27,75	27,75
35	Evan	10	30	50	42,15	27,96	27,75	27,75
36	Stewart	2,75	7	10	6,18	72,36	25,21	6,18
37	Hannah	10	30	59	50,37	19,23	33,48	19,23
38	Victor	8	30	76	66,26	7,82	47,05	7,82
39	Nicola	18	30	75	66,87	3,30	47,68	3,30
40	Olivia	6	7	35	23,71	49,18	0,00	0,00



Gambar 3. Grafik data hasil iterasi 3

Gambar 4. Hasil *Cluster* Iterasi 3

Gambar 4 menunjukkan hasil *cluster* paling optimal yang terbentuk. Dari hasil perhitungan iterasi 3 yang merupakan hasil final dari rangkaian proses K-Means yang sudah dijalankan sebelumnya, dari 40 data yang digunakan diperoleh suatu hasil akhir yang menunjukkan bahwa hasil yang terendah yaitu C1 (*Cluster* 1) dengan kelompok paket data dengan kuota sedikit, lama masa aktif sedang, dan harga paket data terjangkau, hasil sedang yaitu C3 (*Cluster* 3) dengan kelompok paket data dengan kuota sedang, lama masa aktif sedang, dan harga paket data terjangkau, dan hasil yang tertinggi yaitu C2 (*Cluster* 2) dengan kelompok paket data dengan kuota banyak, lama masa aktif sedang, dan harga paket data normal. Berdasarkan hasil uji coba yang diperoleh tersebut menunjukkan bahwa mayoritas penjualan berupa jumlah kuota banyak, masa aktif sedang dan harga normal.

IV. SIMPULAN

Pengelompokan paket penjualan paket data dengan 40 data yang telah diuji dapat dianalisis dengan menggunakan metode K-Means *Clustering*. Berdasarkan uji coba yang telah dilakukan diperoleh hasil optimal pada iterasi ke 3 dengan hasil akhir yang menunjukkan bahwa hasil yang terendah yaitu C1 (*Cluster* 1), hasil sedang yaitu C3 (*Cluster* 3), dan hasil yang tertinggi yaitu C2 (*Cluster* 2). Dari hasil tersebut

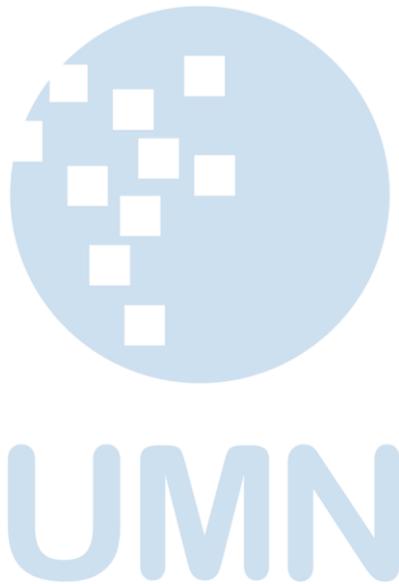
menunjukkan jika penjualan terbanyak berupa kuota banyak menjadi pilihan favorit.

DAFTAR PUSTAKA

- [1] F. Liantoni and L. Cahyani, "Pemanfaatan Hierarchical *Clustering* Untuk Pengelompokan Daun Berdasarkan Fitur Moment Invariant," *J. Ilm. Educat. Univ. Trunojoyo Madura*, vol. 3, no. 2, pp. 91–98, 2017.
- [2] F. Liantoni, N. Ramadijanti, and N. Rosyid Mubtada'i, "Klasifikasi Daun Dengan Centroid Linked *Clustering* Berdasarkan Fitur Bentuk Tepi Daun," *EEPIS Final Proj.*, 2010.
- [3] B. Everitt, S. Landau, M. Leese, and D. Stahl, *Cluster Analysis*. West Sussex, U.K: Wiley: Chichester, 2011.
- [4] A. Al-Wakeel and J. Wu, "K-means Based *Cluster* Analysis of Residential Smart Meter Measurements," *Energy Procedia*, vol. 88, pp. 754–760, Jun. 2016, doi: 10.1016/j.egypro.2016.06.066.
- [5] X. Wu *et al.*, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, Jan. 2008, doi: 10.1007/s10115-007-0114-2.
- [6] Y. Li and H. Wu, "A *Clustering* Method Based on K-Means Algorithm," *Phys. Procedia*, vol. 25, pp. 1104–1109, 2012, doi: 10.1016/j.phpro.2012.03.206.
- [7] Agusta, "K-means Penerapan, Permasalahan dan Metode Terkait," *J. Sist. dan Inform.*, vol. 3, 2007.
- [8] L. Rokach and O. Maimon, "Clustering Methods," in *Data Mining and Knowledge Discovery Handbook*, Springer, Boston, MA, 2005, pp. 321–352.
- [9] K. Tian, J. Li, J. Zeng, A. Evans, and L. Zhang, "Segmentation of tomato leaf images based on adaptive clustering number of K-means algorithm," *Comput.*

[10] P. Fränti and S. Sieranoja, "How much can k-means be

improved by using better initialization and repeats?," *Pattern Recognit.*, vol. 93, pp. 95–112, Sep. 2019, doi: 10.1016/j.patcog.2019.04.014.



Sistem Informasi Klinik Hewan Untuk Meningkatkan Kinerja Keuangan Menggunakan Metode Rasio Profitabilitas

Taufik Kurnialensya

Univ.STEKOM, Fakultas Komputer dan Bisnis, Semarang, Indonesia
taufik@stekom.ac.id

Diterima 27 April 2021

Disetujui 14 Juni 2021

Abstract—Veterinary Clinic XZY Semarang is an agency engaged in serving its clients in caring for the health of livestock and pets. In recording and processing data using an organized system and not yet well integrated. In financial management and also in making financial reports, it is not optimal in evaluating financial performance, so that they cannot see the ability to generate operating profit from the income generated. There is a need for a system to be used to increase activity at the veterinary clinic, so that it is much more effective and efficient in the provision of equipment, equipment, and data management including recording, making financial reports, and assessing financial performance, so that it can show the time used and the system as well. can be accessed simultaneously by interested parties. In the implementation of a new system, namely the Animal Clinic Information System to Improve Financial Performance with the Web-Based Ratio method. From these problems, an application program can be made using php program and MySQL as the database. With the above design, it provides benefits for veterinary clinics that provide benefits in improving their financial performance, can be accessed simultaneously, and provides services in the provision of equipment and equipment needed.

Index Terms—Ratio Method Profitabilitas, Animal Clinic, Information System.

I. PENDAHULUAN

Pengembangan sebuah teknologi memiliki tujuan untuk menaikkan produktivitas dan efektivitas pengguna baik dari SDM maupun organisasi. Perkembangan dalam penggunaan teknologi berbasis web sangat berguna bagi perusahaan atau instansi. Web merupakan sebuah teknologi yang mana pemakai dapat menjalankan secara bersamaan dari satu atau lebih *hardware*, *software*, dan informasi [2].

Perusahaan atau instansi harus mampu berkontribusi didalam membagi informasi, juga bentuk keterangan akan semua proses kegiatan perusahaan yang dikerjakan dalam kurun waktu tertentu, yang mana keterangan itu berkaitan dengan kinerja atau keuangan kepada pimpinan perusahaan[4].

Dalam melakukan sebuah penilaian terhadap kegiatan keuangan dalam perusahaan adalah dengan memakai sebuah analisis rasio probalitas. Penilaian ini biasanya digunakan untuk mengetahui produktifitas dari suatu perusahaan atau instansi [7]. Tingkat profitabilitas perusahaan yang tinggi menunjukkan kemampuan suatu perusahaan atau instansi dalam menciptakan laba atau keuntungan [6].

Klinik Hewan XYZ Semarang merupakan instansi yang bergerak dibidang jasa melayani kliennya dalam merawat kesehatan hewan ternak ataupun hewan peliharaan. Dalam pencatatan dan pengolahan data belum menggunakan sistem yang terorganisasi dan terintegrasi dengan baik, karena dalam masih menggunakan Microsoft Office dalam pengelolaan data.

Dalam sistem ini kurang maksimal pengerjaannya, membutuhkan waktu yang lama dan tingkat kesalahan penginputannya sangat besar, sehingga informasi yang didapatkan tidak akurat atau tidak maksimal. Seringkali bentuk laporan yang disajikan mengalami kesulitan didalam menginterpretasikan terutama yang berkaitan laporan keuangan. Padahal laporan keuangan merupakan manifestasi dari aktivitas yang harus dipertanggung jawabkan [9].

Dalam pengelolaan keuangan dan juga dalam pembuatan laporan keuangan kurang maksimal didalam melakukan evaluasi terhadap kinerja keuangan, sehingga tidak dapat diketahui seberapa besar dan kecil nilai laba pendapatan yang dihasilkan dari sebuah perusahaan. Salah satu cara untuk menilai efisiensi kinerja keuangan dari suatu usaha dalam manajemen keuangan adalah dengan menggunakan analisis rasio profitabilitas. Penilaian profitabilitas ini menggunakan beberapa kriteria antara lain : Net Profit Margin dan Return On Assets.

Net Profit Margin merupakan Rasio yang menggambarkan besarnya laba bersih yang diperoleh oleh perusahaan pada setiap penjualan yang dilakukan, dengan rumus laba bersih dibagi penjualan bersih.

Return On Asset merupakan Rasio yang menggambarkan kemampuan perusahaan untuk menghasilkan keuntungan dari setiap satu rupiah aset

yang digunakan, dengan rumus laba bersih dibagi total aktiva.

Perlu adanya sistem yang digunakan dapat meningkatkan aktivitas pada klinik hewan tersebut, sehingga jauh lebih baik, tepat, dan cermat dalam mempersiapkan perlengkapan, peralatan, dan pengolahan datanya yang meliputi pencatatan, pembuatan laporan keuangan, serta menilai kinerja keuangannya, sehingga dapat meminimalkan waktu yang digunakan dan sistem juga dapat diakses secara bersamaan oleh pihak yang berkepentingan.

Dalam penerapan sebuah sistem yang baru yaitu Sistem Informasi Klinik Hewan Untuk Meningkatkan Kinerja keuangan Menggunakan Metode Rasio Profitabilitas memiliki tujuan sebagai berikut :

1. Menghasilkan sistem informasi klinik hewan yang digunakan untuk menganalisa laporan keuangan sehingga menghasilkan penilaian kinerja keuangan secara valid pada klinik hewan XYZ Semarang.
2. Menciptakan sistem informasi akuntansi pendapatan keuangan yang efektif serta efisien pada klinik hewan XYZ Semarang. Hal tersebut ditandai dengan:
 - a. Memudahkan mengevaluasi kinerja keuangan pada klinik hewan XYZ, sehingga dapat mengetahui kemampuan dalam menghasilkan laba operasi melalui pendapatan operasi yang dihasilkan.
 - b. Aplikasi dapat diakses secara bersamaan oleh Administrasi, Apoteker, dan Pimpinan sehingga dapat dengan mudah mengakses laporan yang dibutuhkan.

Mempermudah dalam penyediaan perlengkapan dan peralatan yang dibutuhkan, serta memberikan kemudahan dalam pengolahan datanya, sehingga dapat meminimalkan waktu yang digunakan

II. TINJAUAN PUSTAKA

A. Sistem Informasi

Sistem informasi adalah sebuah unsur yang memiliki keterkaitan didalam kinerja berdasarkan langkah-langkah yang berhubungan, dan saling mempengaruhi dengan satu dan lainnya dengan tujuan menghasilkan sesuatu.

Komponen sistem informasi diklasifikasikan sebagai berikut : *hardware* dan *software* yang berfungsi sebagai mesin, *people* dan *procedures* yang merupakan manusia dan tatacara menggunakan mesin, Data yang merupakan jembatan penghubung antara manusia dan mesin agar terjadi suatu proses pengolahan data [5].

B. Flowchart

Flowchart merupakan simbol yang memiliki fungsi untuk menjelaskan langkah demi langkah didalam

Flowchart adalah bagan – bagan yang mempunyai arus yang menggambarkan langkah – langkah dalam menyudahi perkara dalam suatu hal. *Flowchart* dibuat menggunakan simbol. Simbol ini digunakan sebagai alat bantu untuk menampilkan proses didalam program. Simbol di dalam *flowchart* dapat dibedakan menjadi dua bagian, yaitu simbol petunjuk dan simbol proses [3].

C. Akuntansi

Akuntansi merupakan bagian dari sistem informasi yang menghitung aktifitas bisnis, mengelola data-data menjadi sebuah laporan, dan menggunakan hasilnya kepada para pengambil keputusan untuk menganalisa hasil dari sebuah laporan. Ditinjau dari sudut kegiatannya, akuntansi dapat didefinisikan sebagai “proses pencatatan, pengolahan, peringkasan, pelaporan, dan penganalisisan data keuangan suatu entitas” [1].

D. Rasio Profitabilitas

Rasio ini mengukur kemampuan para eksekutif perusahaan dalam menciptakan tingkat keuntungan baik dalam bentuk laba perusahaan maupun nilai ekonomis atas penjualan, aset bersih perusahaan maupun modal sendiri (*shareholders equity*). Rasio ini lebih diminati oleh pemegang saham dan manajemen perusahaan sebagai salah satu alat keputusan investasi, apakah investasi bisnis ini akan dikembangkan, dipertahankan, dan sebagainya [1].

E. PHP

PHP (*Hypertext Preprocessor*) adalah salah satu bahasa pemrograman berbasis web, dimana dalam pengopersiannya membutuhkan server yang digunakan untuk mengolah data atau informasi baik itu berupa text, gambar, suara. Bahasa pemrograman PHP ini dibuat secara gratis sehingga orang banyak menggunakan bahasa pemrograman php dibanding dengan yang lainnya [3].

F. MySQL

MySQL merupakan sebuah *database* server yang penggunaannya untuk membuat aplikasi berbasis web, dengan kata lain database ini sebagai induk dalam penyimpanan data atau informasi dan pengolahan datanya. Kepopuleran MySQL antara lain karena MySQL menggunakan SQL (*structure query language*) sebagai bahasa dasar untuk mengakses database-nya sehingga mudah untuk digunakan, kinerja query cepat, serta mencukupi kebutuhan database [3].

III. METODOLOGI PENELITIAN

Borg and Gall dalam “*Educational Research*” (2003) menerangkan : R & D didalam dunia pendidikan merupakan model pengembangan penelitian yang berbasis industri yang mana pengembangan penelitian digunakan dalam mendesain sebuah produk dan dengan langkah baru, setelah itu

dilanjutkan dengan tahapan diuji dilapangan, dievaluasi, dan dilengkapi jika ada kekurangan hingga mendapatkan standar yang telah ditetapkan [8]. Langkah-langkah dalam metodologi penelitian enam langkah :

A. Penelitian Pendahuluan

Langkah pendahuluan ini yaitu melaksanakan dan melakukan pengumpulan data, dengan langkah-langkah seperti :

1. Melakukan wawancara dengan bagian yang berwenang di klinik hewan XYZ Semarang.
2. Melakukan observasi dengan cara mengamati proses kinerja pada klinik hewan XYZ Semarang dalam pembuatan laporan pendapatan.
3. Memahami kajian-kajian yang bersifat literatur dengan tujuan mendapatkan teori atau informasi lainnya sebagai bahan rujukan.

B. Perencanaan

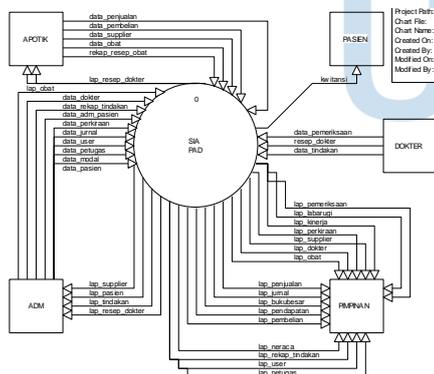
Pada tahap ini dilakukan:

1. Merumuskan tujuan penelitian.
2. Menentukan spesifikasi produk.
3. Membuat jadwal kegiatan.

C. Mengembangkan Bentuk Produk Awal

Membuat perancangan atau model produk dalam bentuk rancang sistem. Hal ini difungsikan supaya didalam pengembangan aplikasi (sistem program) sebagai acuannya. Perancangan ini meliputi *Flow of documant*, *Data Flow Diagram*, *Desain Basis Data*, *Normalisasi*, *Entity Relational Diagram*, tampilan penguana yang berupa rancangan sistem yang akan dibuat.

1. Diagram Konteks



Gambar 1. Konteks Diagram

Keterangan :

- Dari entitas apotik masuk ke sistem berupa data penjualan, data pembelian, data supplier, data obat, dan rekap resep obat. Sedangkan dari sistem menerima laporan resep dokter dan laporan obat.
- Dari entitas pasien menerima laporan dari sistem berupa kwitansi.
- Dari entitas dokter masuk ke sistem berupa resep dokter, data pemeriksaan, dan data tindakan.
- Entitas adm masuk ke sistem berupa data pasien, data dokter, data petugas, data user, data rekap tindakan, data adm pasien, data perkiraan, data modal, dan data jurnal. Sedangkan dari sistem menerima laporan supplier, laporan pasien, laporan tindakan, dan laporan resep dokter.
- Dan pada entitas pimpinan menerima hasil laporan laba rugi, laporan jurnal, laporan buku besar, laporan pendapatan, laporan neraca, laporan kinerja, laporan supplier, laporan perkiraan, laporan dokter, laporan obat, laporan penjualan, laporan pembelian, laporan pemeriksaan, laporan user, laporan petugas, dan laporan rekap tindakan.

2. Normalisasi

Rawatjalan norawat * tglrawat ** nomorm ** nmpemilik jenisewan diaknosa kddokter ** nmdokter kdtindakan ** harga jumlah totrawat nobukti **	Pasien nomorm * tglrekam nmpemilik alamat notelpn jenisewan sex umur beratbadan	Pembelian id_pembelian * nopembelian tglpembelian kdsupplier ** keterangan kdpetugas ** kdobat ** harga jumlah total nobukti **	Penjualan id_penjualan * nopenjualan tglpenjualan nomorm ** ampasien kdobat ** harga jumlah total kdpetugas ** nobukti **	dokter kddokter * nmdokter jeniskelamin tempatlahir tanggalahir alamat notelpn sip	Obat kdobat * nmobat stok keterangan
				Petugas kdpetugas * nmpetugas notelpn jabatan	
				Tindakan kdtindakan * nmtindakan harga keterangan	
		Jurnal id * nobukti ** tanggal keterangan norek ** nama_rekening debit kredit	Supplier kdsupplier * nmsupplier alamat notelpn	Perkiraan norek * nama tipe saldo	

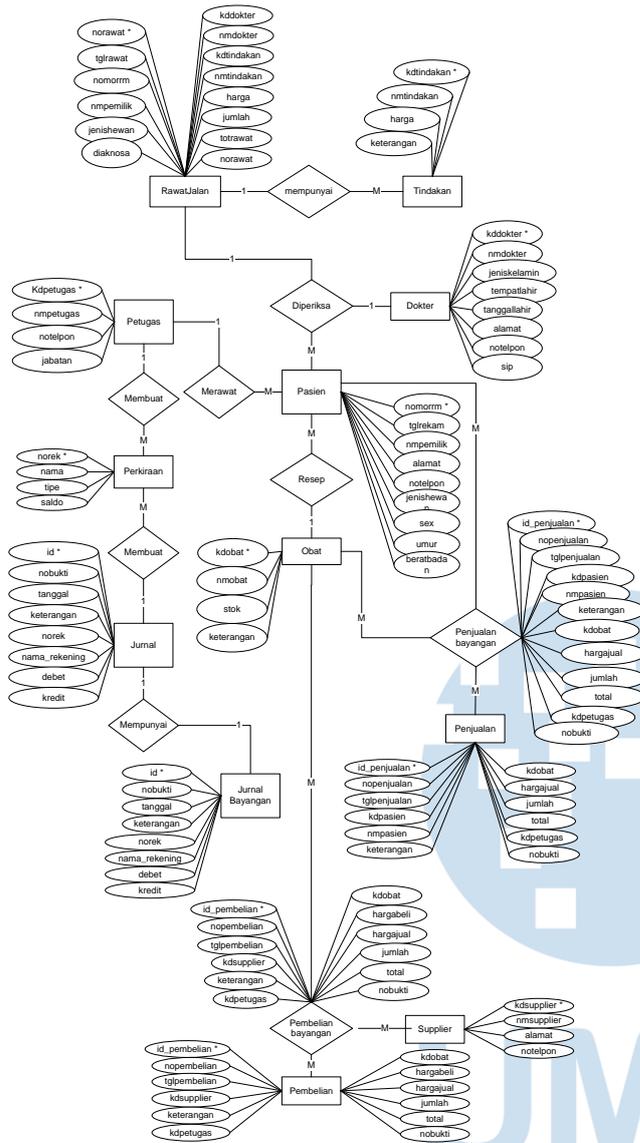
Gambar 2. Normalisasi

Keterangan :

Gambar 2 menjelaskan hubungan antar tabel yang saling mempengaruhi satu dengan lainnya.

Hubungan diantara tabel wajib memiliki atribut kunci primer dan kunci sekunder pada tabel relasinya (* = Primary Key, ** = Secondary Key).

3. ERD



Gambar 3. ERD

4. Basis Data

a. Tabel User

Tabel 1. Struktur Tabel User

Nama Field	Type	Values
id_admin	int	2
nama_admin	varchar	20
username	varchar	20
password	varchar	50
password2	varchar	30
level	enum	(P,AP, ADM)
notelpon	varchar	50

b. Tabel Petugas

Tabel 2. Struktur Tabel Petugas

Nama Field	Type	Values
kdpetugas	varchar	50
nmpetugas	varchar	50
notelpon	varchar	50
jabatan	varchar	50

c. Tabel Rawat Jalan

Tabel 3. Struktur Tabel Rawat Jalan

Nama Field	Type	Values
norawat	varchar	50
tglrawat	date	
nomorm	varchar	50
nmpemilik	varchar	50
jenishewan	varchar	50
diaknosa	varchar	50
kddokter	varchar	50
nmdokter	varchar	50
kdtindakan	varchar	50
nmtindakan	varchar	50
harga	int	11
jumlah	int	11
totrawat	int	11
nobukti	varchar	50

d. Tabel Tindakan

Tabel 4. Struktur Tabel Tindakan

Nama Field	Type	Values
kdtindakan	varchar	50
nmtindakan	varchar	50
harga	int	11
keterangan	varchar	50

e. Tabel Dokter

Tabel 5. Struktur Tabel Dokter

Nama Field	Type	Values
kddokter	varchar	50
nmdokter	varchar	50
jeniskelamin	varchar	50
tempatlahir	varchar	50
tanggallahir	varchar	50
alamat	varchar	50
notelpon	varchar	50
sip	varchar	50

f. Tabel Pasien

Tabel 6. Struktur Tabel Pasien

Nama Field	Type	Values
nomorm	varchar	50
tglrekam	date	
nmpemilik	varchar	50
alamat	varchar	50
notelpon	varchar	50
jenishewan	varchar	50
sex	varchar	50
umur	varchar	50
beratbadan	varchar	50

g. Tabel Perkiraan

Tabel 7. Struktur Tabel Perkiraan

Nama Field	Type	Values
norek	varchar	50
nama	varchar	50
tipe	varchar	50
saldo	int	11

h. Tabel Obat

Tabel 8. Struktur Tabel Obat

Nama Field	Type	Values
norek	varchar	50
nama	varchar	50
tipe	varchar	50
saldo	int	11

i. Tabel Penjualan

Tabel 9. Struktur Tabel Penjualan

Nama Field	Type	Values
id_penjualan	varchar	50
nopenjualan	varchar	50
tglpenjualan	date	
kdpasien	varchar	50
nmpasien	varchar	50
keterangan	varchar	50
kdobat	varchar	50
hargajual	Int	11
jumlah	Int	11
total	Int	11
kdpetugas	varchar	50
nobukti	varchar	50

j. Tabel Penjualanbayangan

Tabel 10. Struktur Tabel Penjualanbayangan

Nama Field	Type	Values
id_penjualan	varchar	50
nopenjualan	varchar	50
tglpenjualan	date	
kdpasien	varchar	50
nmpasien	varchar	50
keterangan	varchar	50
kdobat	varchar	50
hargajual	Int	11
jumlah	Int	11
total	Int	11
kdpetugas	varchar	50
nobukti	varchar	50

k. Tabel Jurnal

Tabel 11. Struktur Tabel Jurnal

Nama Field	Type	Values
id	int	11
nobukti	varchar	50
tanggal	date	
keterangan	text	
norek	varchar	50
nama_rekening	varchar	50
debit	int	11
credit	int	11

l. Tabel Jurnalbayangan

Tabel 11. Struktur Tabel Jurnalbayangan

Nama Field	Type	Values
id	int	11
nobukti	varchar	50
tanggal	date	
keterangan	text	
norek	varchar	50
nama_rekening	varchar	50
debit	int	11
credit	int	11

m. Tabel Pembelian

Tabel 12. Struktur Tabel Pembelian

Nama Field	Type	Values
id_pembelian	Int	11
nopembelian	varchar	50
tglpembelian	date	
kdsupplier	varchar	50
keterangan	varchar	50
kdpetugas	varchar	50
kdobat	varchar	50
hargabeli	Int	11
hargajual	Int	11

jumlah	Int	11
total	Int	11
nobukti	varchar	50

n. Tabel Pembelianbayangan

Tabel 13. Struktur Tabel Pembelianbayangan

Nama Field	Type	Values
id_pembelian	Int	11
nopembelian	varchar	50
tglpembelian	date	
kdsupplier	varchar	50
keterangan	varchar	50
kdpetugas	varchar	50
kdobat	varchar	50
hargabeli	Int	11
hargajual	Int	11
jumlah	Int	11
total	Int	11
nobukti	varchar	50

o. Tabel Supplier

Tabel 14. Struktur Tabel Supplier

Nama Field	Type	Values
Kdsupplier	varchar	50
Nmsupplier	varchar	50
Alamat	varchar	50
Notelpon	varchar	50

D. Melakukan Uji Coba Lapangan

Melakukan uji validasi desain sistem informasi melalui penilaian pengguna (*user*) di klinik hewan XYZ Semarang, desain atau rancangan yang dilakukan penilaian berupa rancang desain form, bentuk laporan, sampai desainnya dinyatakan valid sesuai dengan kebutuhan di klinik hewan XYZ Semarang.

E. Melakukan Revisi Produk Utama

Melakukan perbaikan terhadap desain sistem informasi klinik yang disarankan oleh pengguna (*user*) di klinik hewan XZY Semarang, saat mendapatkan kekurangan dalam pembuatan produk, maka akan dilakukan revisi sesuai kebutuhan yang sudah diminta oleh *user* pengguna sistem.

F. Melakukan Uji Coba Lapangan Utama

Didalam melakukan desain sistem dan dinyatakan sudah sesuai oleh *user* pengguna, maka tahapan selanjutnya adalah membuat *source code* program menjadi *software* aplikasi yang sudah berfungsi dengan baik. Selanjutnya produk *software* akan diuji lapangan oleh calon user di klinik hewan XZY Semarang sampai dinyatakan baik atau efektif dan layak untuk digunakan.

IV. HASIL DAN PEMBAHASAN

Berikut merupakan hasil dari aplikasi yang sudah jadi meliputi :

A. Form Login

Sebelum masuk ke dalam sistem, user yang akan menggunakan aplikasi ini sesuai dengan hak akses yang mereka miliki, maka yang tampil pertama kali

muncul adalah tampilan user login. Didalam form login disini menggunakan level pengguna yaitu terdiri dari pimpinan, Apoteker, dan Administrator (sebagai super admin)

Gambar 4. Form Login

B. Form User

Gambar 5. Form User

Desain form user digunakan untuk memasukkan data pemakai yang terdiri dari nama admin, username, password, level, dan nomor telepon. Dibawah form user terdapat daftar data user yang digunakan untuk menampung data yang akan disimpan.

C. Form Tindakan

Gambar 6. Form Tindakan

Form tindakan digunakan untuk memasukkan data tindakan yang terdiri dari kode tindakan, nama tindakan, harga, dan keterangan. Dibawah form tindakan terdapat daftar data tindakan yang digunakan untuk menampung data yang akan disimpan.

D. Form Petugas

Gambar 7. Form Petugas

Form petugas digunakan untuk memasukkan data petugas yang terdiri dari kode petugas, nama petugas, nomor telpon, dan jabatan. Dibawah form petugas terdapat daftar data petugas yang digunakan untuk menampung data yang akan disimpan.

E. Form Dokter

Gambar 8. Form Dokter

Form dokter digunakan untuk memasukkan data dokter yang terdiri dari kode dokter, nama dokter, jenis kelamin, tempat tanggal lahir, alamat, nomor telpon, dan SIP (surat ijin praktek). Dibawah form dokter terdapat daftar data dokter yang digunakan untuk menampung data yang akan disimpan.

F. Form Obat

Gambar 9. Form Obat

Form obat difungsikan untuk memasukkan data obat yang terdiri dari kode obat, nama obat, stok, dan keterangan. Dibawah form obat terdapat daftar data obat yang digunakan untuk menampung data yang akan disimpan.

G. Form Pasien

No	No Rm	Tanggal	Nama Pemilik	Jenis Hewan	Navigasi
1.	RKM001	2016-09-01	Winda	Kucing	

Gambar 10. Form Pasien

Form pasien digunakan untuk memasukkan data pasien yang terdiri dari nomor rekam medis (No RM), tanggal, nama pemilik, alamat, nomor telpon, jenis hewan, sex (jenis kelamin), umur, dan berat badan. Dibawah form pasien terdapat daftar data pasien yang digunakan untuk menampung data yang akan disimpan.

H. Form Supplier

No	Kode Suppler	Nama Suppler	Alamat	No Telp	Navigasi
1.	SPL001	Wiwini	Semarang	8509090	

Gambar 11. Form Supplier

Form supplier digunakan untuk memasukkan data supplier yang terdiri dari kode supplier, nama supplier, alamat, dan nomor telpon. Dibawah form supplier terdapat daftar data supplier yang digunakan untuk menampung data yang akan disimpan.

I. Form Perkiraan

No	Nomor Rekening	Nama Rekening	Tipe	Saldo	Navigasi
1.	101	Kas	Aktiva	1369500	

Gambar 12. Form Perkiraan

Form perkiraan digunakan untuk memasukkan data perkiraan antara lain nomor rekening, nama rekening, tipe, dan juga saldo. Dibawah form perkiraan terdapat daftar data perkiraan yang digunakan untuk menampung data yang akan disimpan.

J. Form Jurnal

No	No Bukti	Tanggal	Keterangan	No Rekening	Debet	Kredit	Navigasi
1	BEA001	2016-09-09	Pembelian Dari Supplier BEA001	510 - Pembelian	15.000	0	
2	BEA001	2016-09-09	Pembelian Dari Supplier BEA001	101 - Kas	0	15.000	

Gambar 13. Form Jurnal

Form jurnal digunakan untuk memasukkan data jurnal yang terdiri dari nomor bukti, tanggal, keterangan, nomor rekening, nama rekening, debet, dan kredit. Dibawah form jurnal terdapat daftar data jurnal yang digunakan untuk menampung data yang akan disimpan.

K. Form Modal

No	Nomor Rekening	Nama Rekening	Tanggal	Saldo	Navigasi
1.	301	Modal	mm/dd/yyyy		

Gambar 14. Form Modal

Form modal digunakan untuk memasukkan data modal yang terdiri dari nomor rekening, nama rekening, tanggal, dan saldo.

L. Form Pembelian

No	No Pembelian	Tanggal	Kode Obat	Harga Beli	Harga Jual	Jumlah	Total	Navigasi
1.	PEL001	2016-09-08	OBT001	5000	6000	4	20000	

Gambar 15. Form Pembelian

Form pembelian digunakan untuk memasukkan data pembelian yang terdiri dari nomor pembelian, tanggal pembelian, kode supplier, kode petugas, kode obat, harga beli, harga jual, dan jumlah. Dibawah form pembelian terdapat daftar data pembelian yang digunakan untuk menampung data yang akan disimpan. Selain disimpan ke dalam daftar data pembelian, data transaksi pembelian juga tersimpan secara otomatis ke dalam daftar data jurnal.

M. Form Penjualan

Gambar 16. Form Penjualan

Form penjualan digunakan untuk memasukkan data penjualan yang terdiri dari nomor penjualan, tanggal penjualan, pelanggan (No RM), nama pelanggan, kode petugas, kode obat, harga jual, dan jumlah. Dibawah form penjualan terdapat daftar data penjualan yang digunakan untuk menampung data yang akan disimpan. Selain disimpan ke dalam daftar data penjualan, data transaksi penjualan juga tersimpan secara otomatis ke dalam daftar data jurnal.

N. Form Pemeriksaan

Gambar 17. Form Pemeriksaan

Form pemeriksaan digunakan untuk memasukkan data pemeriksaan yang terdiri dari nomor rawat, tanggal rawat, nomor rekam medis (No RM), nama pemilik, jenis hewan, diagnosa, kode dokter, nama dokter, kode tindakan, nama tindakan, harga, dan jumlah. Dibawah form pemeriksaan terdapat daftar data pemeriksaan yang digunakan untuk menampung data yang akan disimpan. Selain disimpan ke dalam daftar data pemeriksaan, data transaksi pemeriksaan juga tersimpan secara otomatis ke dalam daftar data jurnal.

O. Laporan Dokter

No	Kode Dokter	Nama Dokter	TTL	Alamat
1	DTR001	Dth. Heri Gunawan Wibisono	Semarang, 1975-09-06	Semarang
2	DTR002	Dth. Aesa Nomalia	Magelang, 1985-09-15	Magelang

Gambar 18. Laporan Dokter

Laporan dokter digunakan untuk menampilkan data dokter yang terdiri dari kode dokter, nama dokter, tempat tanggal lahir, dan alamat.

P. Laporan Data Obat

No	Kode Obat	Nama Obat	Stok	Keterangan
1	OBT001	Kalosy	21	Antibiotik
2	OBT002	Sulfa Strong	59	Obat Diare
3	OBT003	Ivoamec	50	Obat Kulit
4	OBT004	Gusanex 13 Oz	50	Obat Untuk Luka Kering
5	OBT005	Biosalamin	50	Vitamin Setamina
6	OBT006	Biosan TP	50	Vitamin
7	OBT007	Haemathopan B - 12	51	Suplemen Daya Tahan Tubuh
8	OBT008	Vit B -Flex	50	Vitamin
9	OBT009	Vit B -Komplex	50	Vitamin

Gambar 19. Laporan Data Obat

Laporan data obat digunakan untuk menampilkan data obat yang terdiri dari kode obat, nama obat, stok dan keterangan.

Q. Laporan Data Pasien

No	Kode Pasien	Tanggal RM	Nama Pemilik	Alamat	No Telpun	Jenis Hewan	Sex
1	EKCM001	2016-09-01	Winda	Jl. Atisari Blok Blok H	08190280980	Kucing	Jantan
2	EKCM002	2016-09-10	Ene Della	Jl. Atisari Blok Blok H	081982982738	Unggas	Betina
3	EKCM003	2016-09-10	Kiky	Jl. Mawar No. 15B	08190283423	Anjing	Betina
4	EKCM004	2016-09-12	Ragus	Jl. Pemuda No. 19A	081920839090	Ternak Kecil	Jantan
5	EKCM005	2016-09-12	Vms	Jl. Palang Merah No. 11C	081902808390	Ternak Besar	Betina

Gambar 20. Laporan Data Pasien

Laporan data pasien digunakan untuk menampilkan data pasien yang terdiri dari kode pasien, tanggal RM, nama pemilik, alamat, no telepon, jenis hewan, dan sex.

R. Laporan Data Supplier

No	Kode Supplier	Nama Supplier	Alamat	No Telpun
1	SPL001	Wiwin	Semarang	8509090
2	SPL002	Dewi	Unggaran	8502890
3	SPL003	Suci	Semarang	8502980

Gambar 21. Laporan Data Supplier

Laporan data supplier digunakan untuk menampilkan data supplier yang terdiri dari kode supplier, nama supplier, alamat, dan no telepon.

S. Laporan Data Pembelian Obat

 KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH						
LAPORAN PEMBELIAN OBAT						
No	No Pembelian	Tanggal	Kode Supplier	Kode Obat	Harga Beli	Jumlah
1	PEL002	2016-09-14	SPL002	OBT002	4500	5
2	PEL002	2016-09-14	SPL002	OBT005	5000	5
3	PEL001	2016-09-08	SPL002	OBT001	5000	4
Jumlah Data : 3						

Gambar 22. Laporan Data Pembelian

Laporan data pembelian digunakan untuk menampilkan data pembelian obat. Data yang ditampilkan berupa data no pembelian, tanggal, kode supplier, kode obat, harga beli, dan jumlah.

T. Laporan Data Penjualan Obat

 KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH						
LAPORAN PENJUALAN OBAT						
No	No Penjualan	Tanggal	Nama Pemilik	Kode Obat	Harga Jual	Jumlah
1	PIJ002	2016-09-15	Kiky	OBT001	6000	1
2	PIJ002	2016-09-15	Kiky	OBT002	6000	4
3	PIJ001	2016-09-12	Eno Sella	OBT001	5000	1
Jumlah Data : 3						

Gambar 23. Laporan Data Penjualan Obat

Laporan data penjualan obat digunakan untuk menampilkan data penjualan obat, yang terdiri dari no penjualan, tanggal, nama pemilik, kode obat, harga jual, dan jumlah.

U. Laporan Data Pemeriksaan

 KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH					
LAPORAN PEMERIKSAAN					
No	No Rawat	Tanggal	Nama Pemilik	Jenis Hewan	Dialenosa
1	RW3002	2016-09-16	Bagas	Ternak Kecil	Lemas
2	RW3001	2016-09-14	Kiky	Amjing	Diase
Jumlah Data : 2					

Gambar 24. Laporan Data Pemeriksaan

Laporan data pemeriksaan digunakan untuk menampilkan data pemeriksaan yang terdiri dari no rawat, tanggal, nama pemilih, jenis hewan, dan diagnosa.

V. Laporan Data Jurnal Umum

 KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH						
LAPORAN JURNAL UMUM						
No	No Bukti	Tanggal	Keterangan	No Rekening	Debet	Kredit
1	BEA002	2016-09-08	Pembelian Dari Supplier BEA002	510 - Pembelian	20.000	0
2	BEA002	2016-09-08	Pembelian Dari Supplier BEA002	101 - Kas	0	20.000
3	BEA001	2016-09-09	Pembelian Dari Supplier BEA001	510 - Pembelian	15.000	0
4	BEA001	2016-09-09	Pembelian Dari Supplier BEA001	101 - Kas	0	15.000
Total Transaksi					35.000,00	35.000,00

Gambar 25. Laporan Data Jurnal Umum

Laporan data jurnal umum digunakan untuk menampilkan data jurnal umum yang terdiri dari data nobukti, tanggal, keterangan, no rekening, debet, dan kredit.

W. Laporan Data Buku Besar

 KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH						
LAPORAN BUKU BESAR						
Data Nomor Rekening 101						
No	No Bukti	Tanggal	Keterangan	No Rekening	Debet	Kredit
1	BEA001	2016-09-01	Transaksi Modal	101 - Kas	1.000.000	0
2	BEA002	2016-09-02	Pembelian Dari Supplier BEA002	101 - Kas	0	11.000
3	BEA003	2016-09-05	Transaksi Penjualan Obat Eno Sella	101 - Kas	15.000	0
4	BEA004	2016-09-07	Pembelian Dari Supplier BEA004	101 - Kas	0	15.000
5	BEA005	2016-09-08	Transaksi Penjualan Obat Euge	101 - Kas	16.000	0
6	BEA006	2016-09-09	Pembelian Ramat Jahan	101 - Kas	15.000	0
7	BEA007	2016-09-12	Perolehan	101 - Kas	0	50.000
Saldo					1.844.000,00	76.000,00
Selisih						968.000,00
Data Nomor Rekening 301						
No	No Bukti	Tanggal	Keterangan	No Rekening	Debet	Kredit
8	BEA001	2016-09-01	Transaksi Modal	301 - Modal	0	1.000.000
Saldo					0,00	1.000.000,00
Selisih						1.000.000,00

Gambar 26. Laporan Data Buku Besar

Laporan buku besar digunakan untuk menampilkan data buku besar yang terdiri dari data nobukti, tanggal, keterangan, no rekening, debet, dan kredit.

X. Laporan Data Laba Rugi

 KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH		
LAPORAN LABA RUGI		
No	Pendapatan	Saldo
1	401 - Pendapatan	15000
2	410 - Penjualan	29000
Total Saldo		44.000,00
No	Biaya	Saldo
3	510 - Pembelian	26000
Total Saldo		26.000,00
Laporan Laba Rugi = 18.000,00		

Gambar 27. Laporan Data Laba Rugi

Laporan laba rugi digunakan untuk menampilkan data laba rugi, data yang ditampilkan terdiri dari no rekening, nama rekening dan saldo. Hasil laporan laba rugi didapat dari hasil pendapat dalam satu periode tertentu dikurangi dengan biaya pengeluaran. Jika dalam satu periode tertentu hasil pendapatan lebih besar dari pengeluaran maka dikatakan menerima laba.

Y. Laporan Data Pendapatan

 KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH			
LAPORAN PENDAPATAN			
No	Tanggal	Uraian Rincian Objek	Saldo
1	2016-09-05	410 - Penjualan	13.000
2	2016-09-08	410 - Penjualan	16.000
3	2016-09-09	401 - Pendapatan	15.000
Total Saldo			44.000,00

Gambar 28. Laporan Data Pendapatan

Laporan data pendapan digunakan untuk menampilkan data pendapatan yang terdiri dari data tanggal, uraian pendapatan, dan saldo.

Z. Laporan Data Net Profit Margin (NPM)

KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH			
Periode Pertama Tanggal 2016-08-01 sampai dengan 2016-08-31		Periode Kedua Tanggal 2016-09-01 sampai dengan 2016-09-30	
Analisa NPM	Saldo	Analisa NPM	Saldo
Laba Kotor	920,000	Laba Kotor	1,572,000
Beban Usaha	530,000	Beban Usaha	20,000
Laba Bersih	390,000	Laba Bersih	1,552,000
Penjualan Bersih	20,000	Penjualan Bersih	69,000
NPM	16.58	NPM	21.99

Rata - Rata Industri NPM = 20%
 Jika : NPM >= 20% = BAIK
 NPM <= 20% = KURANG BAIK

ANALISA

Periode Pertama Tanggal 2016-08-01 sampai 2016-08-31 :
NPM : 16.58 (KURANG BAIK)

Periode Kedua Tanggal 2016-09-01 sampai 2016-09-30 :
NPM : 21.99 (BAIK)

Jadi, Kinerja pada Klinik Hewan **MENGALAMI KENAIKAN**

Gambar 29. Laporan Data NPM

Laporan data *net profit margin* digunakan untuk menampilkan laporan keuangan NPM. Bentuk rumus NPM adalah Laba Bersih dibagi Penjualan Bersih. Dari klasifikasi rata-rata industri NPM 20%. Jika NPM >=20% maka tergolong baik. Jika NPM <=20% maka tergolong kurang baik.

Hasil analisa pada periode tanggal 01-08-2016 sampai 31-08-2016 hasil NPM sebesar 16.58 (kurang baik). Pada periode tanggal 01-09-2016 sampai 30-09-2016 nilai NPM sebesar 21.99 (baik), maka terdapat kenaikan nilai NPM pada klinik hewan XYZ.

AA. Laporan Data Return On Asset (ROA)

KLINIK HEWAN XZY SEMARANG PROPINSI JAWA TENGAH			
Periode Pertama Tanggal 2016-08-01 sampai dengan 2016-08-31		Periode Kedua Tanggal 2016-09-01 sampai dengan 2016-09-30	
Analisa ROA	Saldo	Analisa ROA	Saldo
Laba Kotor	920,000	Laba Kotor	5,572,000
Beban Usaha	530,000	Beban Usaha	770,000
Laba Bersih	390,000	Laba Bersih	4,081,700
Total Aktiva	1,890,000	Total Aktiva	4,802,000
ROA	0.18	ROA	0.85

Rata - Rata Industri ROA = 30%
 Jika : ROA >= 30% = BAIK
 ROA <= 30% = KURANG BAIK

ANALISA

Periode Pertama Tanggal 2016-08-01 sampai 2016-08-31 :
ROA : 0.18 (KURANG BAIK)

Periode Kedua Tanggal 2016-09-01 sampai 2016-09-30 :
ROA : 0.85 (KURANG BAIK)

Jadi, Kinerja pada Klinik Hewan **MENGALAMI KENAIKAN**

Gambar 30. Laporan Data ROA

Laporan data *return on assest* digunakan untuk menampilkan laporan keuangan ROA. Bentuk rumus ROA adalah laba bersih dibagi total aktiva. Untuk klasifikasi ROA memiliki standart jika nilai ROA >=30% maka kategori baik, jika <= 30% kategori kurang baik. Dari hasil analisa pada gambar diatas pada periode 01-08-2016 sampai 30-08-2016 nilai ROA sebesar 0.18 sehingga kategori kurang baik.

Pada periode 01-09-2016 sampai 30-09-2016 nilai ROA sebesar 0.85 kategori kurang baik.

V. SIMPULAN DAN SARAN

Dari pemaparan, maka dapat diambil kesimpulan sebagai berikut :

1. Pengelolaan data keuangan pada klinik hewan xyz yang sebelumnya masih bersifat tunggal, sehingga mengalami kesulitan dalam penyampaian informasi, data tidak dapat diakses secara bersamaan, penyusunan data juga belum tersusun dengan baik, dan juga belum adanya penilaian kinerja keuangan pada klinik hewan XYZ tersebut. Sehingga sistem yang digunakan pada klinik hewan belum efektif.

Dari permasalahan tersebut mampu dibuatkan aplikasi sistem informasi klinik hewan untuk meningkatkan kinerja keuangan dengan menggunakan metode rasio profitabilitas dengan menggunakan bahasa pemrograman php dan MySQL sebagai databases-nya. Sehingga memberikan manfaat bagi klinik hewan xyz seperti memberikan kemudahan dalam mengevaluasi kinerja keuangannya, aplikasi dapat diakses secara bersamaan, dan memberikan kemudahan dalam penyediaan perlengkapan dan peralatan yang dibutuhkan.

2. Hasil penilaian terhadap produk dari ahli pakar adalah sebagai berikut :

Tabel 15 Hasil Penilaian Rancang Desain Program

No	INDIKATOR	1	2	3	4
1	Apakah rancang desain DFD (Data Flow Diagram) sudah tepat ?				√
2	Apakah rancang desain antarmuka (Interface) untuk sistem sudah tepat ?			√	
3	Apakah rancang desain normalisasi data sudah tepat ?			√	
4	Apakah rancang desain ERD (Entitiy Relational Diagram) sudah tepat ?			√	
5	Apakah rancang desain struktur tabel basis data sudah tepat ?			√	
6	Apakah rancang desain input program sudah tepat ?			√	
7	Apakah rancang desain output program sudah tepat ?			√	
8	Apakah rancang desain tampilan laporan program sudah tepat ?			√	
9	Apakah rancang desain menu – menu pada program sudah sesuai kebutuhan ?			√	
10	Apakah rancang desain program mudah dipahami oleh pengguna ?			√	
JUMLAH SKOR		31			
TOTAL JUMLAH SKOR		31			

Tabel 16 Skala Kriteria Validasi

Nilai	Kriteria Kevalidan
3,26 – 4,00	Sangat Valid
2,51 – 3,25	Valid (Baik)
1,76 – 2,50	Kurang Valid (Revisi)
1,00 – 1,75	Tidak Valid (Revisi total)

Tabel 17 Hasil Pengisian Angket User

NO	INDIKATOR	User 1	User 2	User 3
1	Kenyamanan melihat tampilan muka dan lembar – lembar berikutnya	3	4	3
2	Adakah kemudahan melakukan input data keuangan	4	3	3
3	Apakah informasi – informasi tampilan memberikan kemudahan pengguna	3	3	4
4	Apakah pengguna merasa senang mengakses karena cepat memperoleh informasi	3	3	4
5	Apakah pengguna merasa puas karena informasi yang ada sesuai kebutuhan pengguna	3	4	3
6	Pengguna akan terus mengakses karena dimungkinkan proses cepat	3	4	3
7	Apakah output sudah sesuai dengan harapan ?	4	4	3
8	Apakah pengguna merasa program ini sudah menyajikan laporan keuangan yang memperlihatkan posisi kinerja keuangan	3	3	3
9	Apakah pengguna merasakan perubahan antara sistem lama dengan yang baru	4	3	4
10	Apakah hasil pemilihan SIA ini mempermudah dalam pengambilan keputusan	4	3	4
JUMLAH SKOR		34	34	34
TOTAL JUMLAH SKOR		102		

Jumlah total nilai sebanyak 3 responden adalah 102, maka dapat dihitung rata-rata = total Nilai responden/ jumlah rponden ($102/3 = 34$). Maka diperoleh nilai validasi dari responden sebagai berikut $34/ 10 = 3.4$

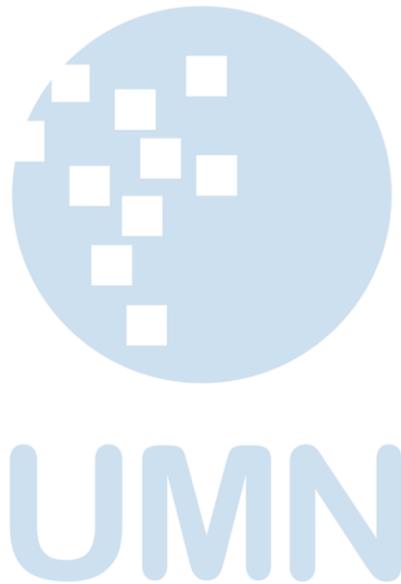
Dapat diambil kesimpulan bahwa nilai yang diperoleh berada diantara 2.51 sampai 3.25 kategori valid atau baik. Dari hasil penilaian validator produk dari user atau pengguna dapat disimpulkan bahwa nilai yang diperoleh 3.4 berada diantara 3.26 sampai 4.00 kategori sangat valid.

Setelah pembuatan sistem informasi klinik hewan untuk meningkatkan kinerja keuangan ini dilakukan, ada beberapa saran yang perlu dikembangkan pada penelitian selanjutnya, yaitu :

1. Pada sistem informasi klinik hewan untuk meningkatkan kinerja keuangan dengan metode rasio profitabilitas keuangan masih bersifat *offline*. Diharapkan untuk pengembangan selanjutnya memberikan sistem informasi klinik hewan yang dapat diakses secara *online*.
2. Pada sistem informasi klinik hewan ini hanya menerapkan metode rasio profitabilitas. Diharapkan untuk pengembangan selanjutnya memberikan sistem informasi akuntansi penilaian kinerja keuangan dengan menerapkan metode rasio keuangan yang lebih lengkap dengan cara menambahkan komponen lain yang bisa dimasukkan ke dalam produk sehingga laporan kinerja keuangan yang dihasilkan lebih lengkap.

DAFTAR PUSTAKA

- [1] Arfan Ikhsan Lubis ,” Akuntansi Keperilakuan: Akuntansi Multiparadigma” 2018, Salemba Empat Jakarta.
- [2] Dahlia MO, Mahdiana, “Penggabungan Metode Analytical Hierarchy Process Dan Multi-Objective Optimization On The Basis Of Ratio Analysis Untuk Pemilihan Pegawai Terbaik Pada Klinik Yadika Petukangan,” Jurnal IDEALIS, Vol 3 No.1 , 2020, hal 123-132.
- [3] Dewi, K.Soedarsono, Dkk, Sistem Manajemen Komunikasi (Teori, Model, dan Aplikasi) Simbiosis Rekatama Media. 2020
- [4] Isti Aning Ika Haryani,” Rancang Bangun Sistem Informasi Keuangan Menggunakan Metode Analisis Rasio keuangan Berbasis Website,” JMI, Vol 10, No.1, 2019, hal 49-57.
- [5] James A Brien, Geoge M. Marakas “Pengantar Sistem Informasi” Salemba Empat, 2017 Gudang Penerbit.
- [6] Leny Mirta Sari, Achmad Arwan, Djoko Pramono,” Pengembangan Sistem Informasi Rekam Medis Hewan dan Rekapitulasi Transaksi Berbasis Web (Studi Kasus : Klinik Hewan House Of Pet Malang),” JPTIHK Vol 4, No.11, 2020, hal 3905-3912.
- [7] Nadia Pranata, Tiara Nur Safitri,” Sistem Pendukung Keputusan Menentukan Kelanjutan Dan Kinerja Keuangan Pada Usaha Produksi Gas Menggunakan Metode Rasio Likuiditas,” JSON, Vol 2 No.2, 2020, hal 206-211.
- [8] Nusa Putra, 2012,” Research dan Development/ Penelitian dan Pengembangan : Suatu Pengantar /Nusa Putra,” Raja Grafindo : Jakarta.



Perbandingan *Convolutional Neural Network* pada *Transfer Learning Method* untuk Mengklasifikasikan Sel Darah Putih

Daniel Martomanggolo Wonohadidjojo

Program Studi Informatika, Fakultas Teknologi Informasi, Universitas Ciputra Surabaya, Surabaya, Indonesia
daniel.m.w@ciputra.ac.id

Diterima 17 Mei 2021

Disetujui 13 Juni 2021

Abstract—Analysis of White Blood Cells structure from microscopic images and classification of cells into types is challenging. Although white blood cells can be differentiated based on their shape, color and size, one challenging aspect is that they are surrounded by other blood components such as red blood cells and platelets. In this study, *transfer learning* method using four *network* architectures that have been trained in advance is applied to classify the white blood cell images. The *network* architectures used are AlexNet, GoogleNet, ResNet-50 and Visual Geometry Group (VGG) - 16. A comparative analysis of the performance of these architectures was carried out in classifying the images. The evaluation method was undertaken using *Confusion matrix*. The *performance metrics* measured in the evaluation are *Accuracy*, *Precision*, *Recall* and *Fmeasure*. The results showed that all architectures succeeded in classifying white blood cells using the *transfer learning* method. ResNet-50 is the *network* architecture that shows the highest performance in classifying white blood cell images.

Index Terms—white blood cells, *transfer learning*, AlexNet, GoogleNet, ResNet-50, VGG-16

I. PENDAHULUAN

Darah terutama mengandung tiga jenis sel: sel darah merah, sel darah putih dan trombosit. Dari tiga jenis tersebut, sel darah putih adalah jenis sel darah yang dibuat di sumsum tulang dan ditemukan di darah dan jaringan getah bening. Sel darah putih merupakan bagian dari sistem kekebalan tubuh. Mereka membantu tubuh melawan infeksi dan penyakit lainnya. Sel darah putih dapat dikategorikan menjadi dua jenis, ditentukan oleh penampilan sitoplasma. Jenis pertama adalah Granulosit dan termasuk Basophil, *Eosinophil* dan *Neutrophil*. Kelompok kedua, yang disebut Agranulosit, termasuk *Lymphocyte* dan *Monocyte* [1].

Sel darah putih memiliki fungsi penting untuk sistem kekebalan tubuh, karena sel darah putih merupakan pertahanan utama tubuh terhadap infeksi dan penyakit. Memeriksa jumlah sel darah putih dalam darah biasanya merupakan bagian dari tes sel darah lengkap. Ini dapat digunakan untuk mencari kondisi

seperti infeksi, peradangan, alergi, dan leukemia. Hal ini merupakan masalah tersendiri karena kategori sel darah putih mencakup sel-sel yang dapat sangat bervariasi satu sama lain. Meskipun sel darah putih dapat dibedakan berdasarkan bentuk dan ukurannya, satu aspek yang menantang adalah sel darah putih dikelilingi oleh komponen darah lain seperti sel darah merah dan trombosit.

Berdasarkan penjelasan tersebut, maka diperlukan metode untuk mengklasifikasikan sel darah putih secara akurat.

Deep Learning menggunakan *Convolution Neural Networks* (CNN) [2] saat ini merupakan pilihan terbaik dalam aplikasi pencitraan medis seperti deteksi dan klasifikasi [3]. Meskipun CNN mencapai hasil terbaik pada kumpulan data yang besar, tapi CNN membutuhkan banyak data dan sumber daya komputasi untuk dilatih. Dalam banyak kasus, kumpulan data terbatas dan mungkin tidak cukup untuk melatih CNN dari awal. Dalam skenario seperti itu, untuk memanfaatkan kekuatan CNN dan pada saat yang sama mengurangi sumber daya komputasi, *transfer learning* dapat digunakan [4]. Dalam pendekatan ini, CNN awalnya dilatih pada kumpulan data citra umum yang besar dan beragam dan kemudian diterapkan ke tugas tertentu.

Terdapat beberapa *neural network* pra-terlatih yang telah memenangkan kompetisi internasional seperti VGGNet [5], ResNet [6], GoogleNet [7], dan AlexNet [8]. Dalam [9] dilakukan evaluasi arsitektur CNN yang berbeda, dan *transfer learning* mencapai kinerja skor tertinggi pada kelenjar getah bening torako-abdominal serta klasifikasi penyakit paru interstitial.

Beberapa studi telah dilakukan untuk menerapkan CNN pada klasifikasi citra sel darah putih. Ekstraksi fitur dari citra sel darah putih menggunakan VGGNet, yang kemudian difilter menggunakan *Salp Swarm Algorithm* (SESSA) yang ditingkatkan secara statistik dilakukan pada [1]. Pada [10] diterapkan klasifikasi CNN untuk mengeksplorasi kelayakan pendekatan *deep learning* untuk mengidentifikasi limfosit dan

subtipe acute *lymphoblastic leukaemia*, dan dibandingkan dengan pendekatan *Support Vector Machines* (SVM). Selain itu, dua klasifikasi pembelajaran mesin tradisional, *multilayer perceptron* (MLP), dan *random forest* juga diterapkan untuk perbandingan.

Suatu model *Deep Learning* berbasis *Hybrid CNN* untuk mengklasifikasikan leukosit eosinofil, limfosit, monosit, dan *neutrophil* diusulkan pada [11]. Model yang disajikan tersebut didasarkan pada arsitektur AlexNet dan GoogleNet yang telah dilatih sebelumnya. Fitur Vektor pada *pooling layer* terakhir pada kedua arsitektur CNN telah digabungkan, dan kemudian fitur vektor yang dihasilkan diklasifikasikan oleh SVM.

Pada [12][13] dilakukan kombinasi beberapa arsitektur *deep learning* untuk meningkatkan kegunaan *transfer learning* dalam mengklasifikasikan citra berbasis sel.

Berdasarkan seluruh penjelasan tersebut maka pada penelitian ini digunakan *transfer learning* untuk melatih ulang CNN untuk mengklasifikasikan citra sel darah putih.

Selanjutnya paper ini disusun sebagai berikut: section II menjelaskan tentang kajian Pustaka yang digunakan pada penelitian ini, section III menjelaskan tentang metode penelitian yang digunakan, section IV menjelaskan tentang hasil uji coba dan eksperimen serta pembahasannya dan section V menjelaskan tentang kesimpulan penelitian ini.

II. KAJIAN PUSTAKA

A. Transfer Learning

CNN yang digunakan untuk penerapan *transfer learning*, sebelumnya telah di-*training* dengan menggunakan dataset citra sejumlah obyek/ benda. Pada *training* tersebut, *network* dapat mengklasifikasikan citra ke dalam sejumlah kategori obyek, seperti *keyboard*, cangkir kopi, pensil, dan banyak hewan. Dengan demikian *network* tersebut telah mempelajari representasi fitur yang kaya untuk berbagai citra. Selanjutnya *network* mengambil citra sebagai input dan kemudian mengeluarkan label untuk tiap obyek dalam citra bersama dengan probabilitas untuk masing-masing kategori obyek.

Transfer learning merupakan metode yang menggunakan *network* yang sudah dilatih sebelumnya dan menggunakannya sebagai titik awal untuk mempelajari tugas baru. Melakukan fine tuning *network* dengan *transfer learning* jauh lebih cepat dan lebih mudah daripada melatih *network* dari awal dengan bobot yang diinisialisasi secara acak.

Transfer learning ini dapat mentransfer fitur yang dipelajari ke tugas baru dengan menggunakan lebih sedikit dataset pada proses *training* sehingga prosesnya lebih cepat. Melakukan *fine tuning network*

biasanya lebih lambat dan membutuhkan lebih banyak upaya daripada ekstraksi fitur sederhana. Tetapi jika digunakan *transfer learning* maka karena *network* sudah di-*training* sebelumnya dan mempelajarinya untuk mengekstrak serangkaian fitur yang berbeda, maka *network* yang baru menjadi lebih akurat. Dengan demikian maka fine-tuning pada *network* yang baru dapat bekerja lebih baik daripada ekstraksi fitur meskipun hanya menggunakan dataset dalam jumlah yang lebih kecil, karena *network* yang baru tersebut telah belajar untuk mempelajari fitur yang baru.

B. Evaluasi performa

Untuk mengukur performa model *network* tersebut dalam melakukan klasifikasi digunakan *Confusion matrix* (CM). Pada CM nilai prediksi dibandingkan dengan nilai aktual. Nilai prediksi adalah keluaran dari program dimana nilainya Positif dan Negatif sedangkan nilai aktual adalah nilai sebenarnya yang berupa nilai True dan False. Dengan demikian terdapat 4 kemungkinan hasil prediksi yaitu True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). Berdasarkan kemungkinan tersebut maka *performance metrics* yang digunakan direpresentasikan pada persamaan (1) – (5) [14][15].

Accuracy merupakan *metric* yang banyak digunakan pada evaluasi performa tersebut. *Metric* ini menghitung persentase sampel yang diklasifikasikan dengan benar, dan direpresentasikan pada Persamaan (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (1)$$

Precision mengukur seberapa "tepat" model tersebut berdasarkan pada yang diprediksi positif dan berapa banyak di antaranya yang sebenarnya positif. *Metric* ini dihitung menggunakan persamaan (2).

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Sensitivity atau *Recall* merupakan *metric* yang mengukur kesesuaian dengan keakuratan contoh positif, dan mengacu pada berapa banyaknya contoh kelas positif diberi label dengan benar. *Metric* ini dihitung dengan persamaan (3), di mana TP mengacu pada true positive, yang merupakan jumlah kejadian yang positif dan diidentifikasi dengan benar, dan FN mewakili *false negative*, yang merupakan jumlah kasus positif yang salah diklasifikasikan sebagai negatif.

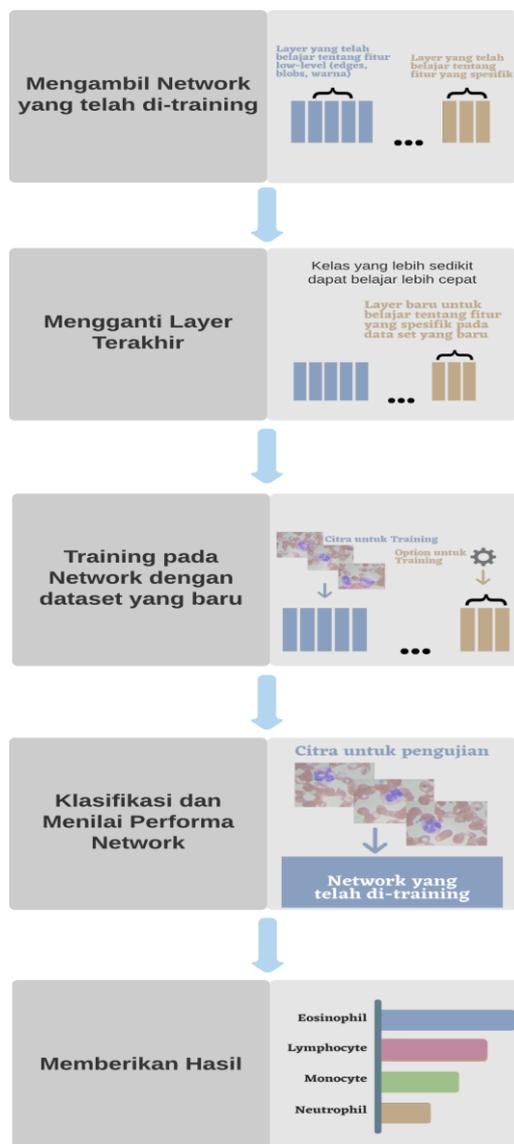
$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Fmeasure merupakan perbandingan rata-rata presisi dan *recall* yang dibobotkan. *Metric* ini dihitung menggunakan persamaan (4).

$$F \text{ measure} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

III. METODE PENELITIAN

Network terlatih yang digunakan pada *transfer learning* ini adalah model CNN. CNN sama dengan *neural network* lainnya, terdiri dari fungsi bobot, bias, dan aktivasi. CNN memiliki 2 bagian besar *layer* yaitu *layer* untuk *feature extraction* dan *layer* untuk klasifikasi. *Layer* untuk *feature extraction* terdiri dari *layer convolutional*, *pooling layer*, *stride*, dan *padding*. Sedangkan *layer* untuk klasifikasi terdiri dari *fully connected layer*, *softmax*, dan *layer* keluaran. *Network* terlatih inilah yang digunakan pada metode *transfer learning* pada penelitian ini. Tahap *transfer learning* yang digunakan pada penelitian ini meliputi: meng-*import network* yang sudah dilatih sebelumnya, mengganti *layer* klasifikasi, melatih *network* pada data citra sel darah putih, mengklasifikasikan sel darah putih dan mendapatkan ukuran kinerja yang akurat. Gambar 1 menunjukkan metode *transfer learning* yang digunakan pada penelitian ini.



Gambar 1. Metode *Transfer learning* yang digunakan

A. Meng-import Network Yang Telah Dilatih

Pada penelitian ini arsitektur *network* yang digunakan adalah AlexNet, GoogleNet, ResNet-50, dan VGG-16 sebagai *network* yang telah dilatih. Setiap *network* terlatih memiliki konfigurasi *layer* yang berbeda. *Layer* di awal dan tengah disebut *layer feature extraction*. *Layer* tersebut menghasilkan fitur sederhana seperti *brightness* dan *edge*, hingga fitur kompleks yang unik seperti warna dan bentuk. Hasil pembelajaran pada *layer feature extraction* pada lingkungan sumber (lingkungan dimana *network* tersebut berasal) dapat ditransfer ke *layer feature extraction* pada lingkungan target (lingkungan baru pada penelitian ini). Pembelajaran *feature extraction* pada lingkungan target digunakan untuk meng-ekstrak citra data sel darah putih yang digunakan pada tahap *training*.

Arsitektur GoogLeNet memiliki 22 *layer* jika dihitung hanya *layer* yang memiliki parameter (atau 27 lapisan jika kita juga menghitung *pooling*). *Network* ini pada lingkungan asalnya telah dilatih pada lebih dari satu juta citra dan dapat mengklasifikasikan citra ke dalam 1000 kategori obyek. Arsitektur *network* ini telah mempelajari representasi fitur yang kaya untuk berbagai citra. *Network* mengambil citra sebagai input dan output label untuk obyek dalam citra bersama-sama dengan probabilitas untuk setiap kategori obyek [7].

Arsitektur AlexNet berisi delapan *layer* dimana lima yang pertama adalah *layer convolutional*, beberapa di antaranya diikuti oleh *layer max-pooling*, dan tiga yang terakhir adalah *layer fully-connected*. Arsitektur *network* ini menggunakan fungsi aktivasi ReLU non-jenuh, yang menunjukkan peningkatan kinerja *training* di atas tanh dan sigmoid [8].

Arsitektur ResNet-50 adalah Jaringan *Residual* yang terdiri dari 50 *layer*. Seperti yang ditunjukkan oleh namanya, *network* ini menggunakan pembelajaran *Residual*. Pada pembelajaran *Residual*, *network* bukannya mencoba mempelajari beberapa fitur, tetapi *network* ini mempelajari beberapa *Residual*. *Residual* secara sederhana dapat dipahami sebagai pengurangan fitur yang dipelajari dari masukan suatu *layer*. ResNet melakukan ini dengan menggunakan koneksi pintasan yaitu langsung menghubungkan input dari *layer* ke- n ke beberapa *layer* berikutnya ($n + x$). Telah dibuktikan bahwa melatih bentuk *network* ini lebih mudah daripada melatih *deep convolutional neural networks* yang sederhana dan juga masalah penurunan akurasi dapat teratasi [6].

VGG16 adalah arsitektur CNN yang digunakan untuk memenangkan kompetisi ILSVR (Imagenet) pada tahun 2014. Ini dianggap sebagai salah satu arsitektur model vision yang sangat baik hingga saat ini. Hal yang paling unik tentang VGG16 adalah bahwa alih-alih memiliki sejumlah besar *hyper-*

parameter, network ini memiliki *layer convolution* dengan filter 3x3 dan *layer maxpool* dengan filter 2x2. Konfigurasi *layer convolution* dan *maxpool* ini diterapkan secara konsisten di seluruh arsitektur. *Network* ini memiliki 2 *layer fully connected* yang diikuti oleh *softmax* untuk output. Arsitektur VGG-16 memiliki 16 *layer* yang memiliki bobot. *Network* ini adalah jaringan yang cukup besar dan memiliki sekitar 138 juta parameter [5].

B. Transfer learning

Pada penelitian ini digunakan metode *Transfer learning* untuk pembelajaran pada *network* yang memanfaatkan *layer classification*. *Layer classification* dikenal sebagai 3 *layer* terakhir yaitu *layer fully connected*, *softmax*, dan *output*. Tabel 1 menunjukkan nama-nama *layer* yang digunakan pada empat arsitektur *network* yang digunakan pada penelitian ini.

Tabel 1. Classification *layer* pada *network* yang telah dilatih

Network yang telah dilatih	Classification Layer
AlexNet	fc8, prob, output
GoogleNet	loss3-classifier, prob, output
ResNet50	fc1000, fc1000_softmax, classificationlayer_fc1000
VGG16	fc8, prob, output

Layer classification yang semula ini diganti dengan *layer classification* baru yang sesuai dengan tugas klasifikasi baru. *Layer* yang baru ini juga mencakup jumlah kelas baru dan kecepatan pembelajaran yang ditetapkan di *network* yang baru.

Layer classification pada *network* asalnya dikonfigurasi untuk jumlah kelas pada *training* sebelumnya. *Layer* ini pada arsitektur *network* tersebut berisi informasi yang diproses oleh *network* menjadi probabilitas kelas, nilai loss, dan label yang diprediksi. Untuk mengklasifikasikan citra baru, *layer* ini diganti dengan *layer* baru dan disesuaikan dengan dataset yang baru.

Layer fully connected diganti dengan *layer* yang baru yang memiliki jumlah keluaran yang sama dengan jumlah kelas. Untuk membuat pembelajaran lebih cepat pada *layer* yang baru dibandingkan dengan *layer* sebelumnya, faktor kecepatan pembelajaran ditingkatkan.

Layer classification menentukan kelas keluaran *network*. *Layer classification* diganti dengan yang baru tanpa label kelas. Pada saat *training*, *network* secara otomatis menentukan kelas keluaran dari *layer* tersebut.

Network memerlukan citra masukan dengan ukuran tertentu, tetapi citra input memiliki ukuran yang berbeda. Ukuran citra diubah secara otomatis

untuk *training network* tersebut. Selanjutnya dilakukan operasi augmentasi data pada citra *training* dengan menerapkan transformasi geometris affine acak yaitu teknik *reflection* secara horizontal, dan teknik *translation* secara horizontal dan vertikal. Augmentasi data membantu mencegah *network* dari *overfitting* dan mengingat detail yang tepat dari citra *training*.

Pada *training* ditentukan parameter kecepatan pembelajaran untuk *transfer learning* dengan mempertahankan fitur pada *layer* awal *network* yang telah dilatih sebelumnya (bobot *layer* yang ditransfer). Untuk memperlambat pembelajaran di *layer* yang ditransfer, kecepatan pembelajaran awal di-set ke nilai yang kecil. Pada langkah sebelumnya, telah dilakukan peningkatan faktor kecepatan pembelajaran untuk *layer fully-connected* untuk mempercepat pembelajaran di *layer* akhir yang baru. Kombinasi *setting* kecepatan pembelajaran ini menghasilkan pembelajaran yang berjalan cepat hanya di *layer* baru dan pembelajaran lebih lambat di *layer* yang lain.

Saat melakukan *transfer learning*, tidak perlu dilakukan *training* dengan jumlah *epoch* yang sama dengan jumlah sebelumnya. *Epoch* adalah siklus *training* lengkap di seluruh dataset *training*. Pada *training* ini juga ditentukan ukuran *mini batch file* dan data validasi. Validasi *network* dilakukan setiap iterasi dengan frekuensi yang telah ditentukan selama *training*.

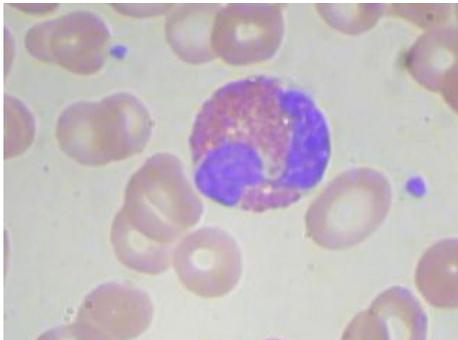
Selanjutnya *training* dilakukan pada *network* yang terdiri dari *layer* yang ditransfer dan *layer* baru dengan menggunakan GPU. Kemudian klasifikasi citra dilakukan dengan menggunakan *network* yang telah melalui *training* tersebut.

C. Dataset dan eksperimen

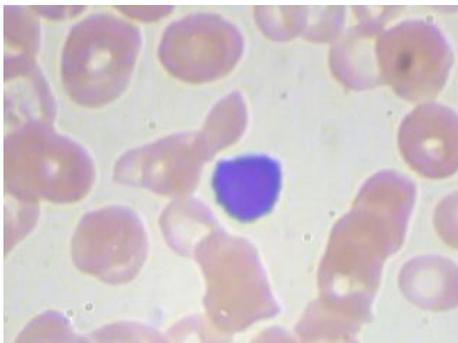
Dataset untuk *training* didapatkan dari Blood Cells Image yang tersedia secara publik pada <https://www.kaggle.com/paultimothymooney/blood-cells>. Pada *training* digunakan 400 citra digital sel darah putih yang terdiri dari 4 kelas yaitu 100 citra *Eosinophil*, 100 citra *Lymphocyte*, 100 citra *Monocyte* dan 100 citra *Neutrophil*. Dengan jumlah dataset untuk *training* sebanyak 400 ini maka proses *training* dapat berlangsung lebih cepat dibandingkan dengan menggunakan jumlah dataset lebih banyak dalam jumlah ribuan. Dengan jumlah dataset tersebut fine-tuning pada *network* yang baru tetap dapat bekerja lebih baik, karena *network* yang baru itu telah belajar untuk mempelajari fitur yang baru tersebut sebagai hasil *training* sebelumnya. Hal ini merupakan kelebihan metode *transfer learning* yang digunakan dalam penelitian ini.

Ukuran tiap citra digital adalah piksel. Pengaturan eksperimental adalah sebagai berikut: dataset dibagi menjadi dua bagian, 70% untuk *training* dan 30% untuk pengujian. Total data untuk tiap kelas adalah

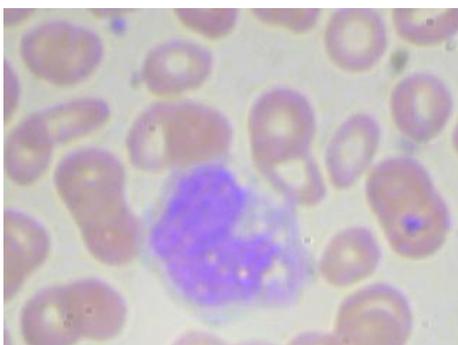
100 citra, 70 citra untuk *training*, dan 30 citra untuk validasi. Ukuran semua citra adalah 320x240 piksel. Untuk keperluan *training* ini dimana diperlukan variasi citra, maka juga digunakan citra yang diaugmentasi. Gambar 2, Gambar 3, Gambar 4 dan Gambar 5 berturut-turut menunjukkan contoh citra *Eosinophil*, *Lymphocyte*, *Monocyte* dan *Neutrophil*.



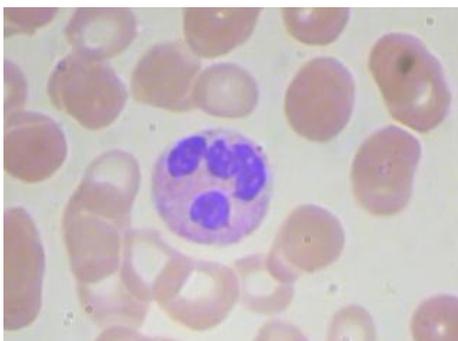
Gambar 2. Citra *Eosinophil*.



Gambar 3. Citra *Lymphocyte*.



Gambar 4. Citra *Monocyte*.



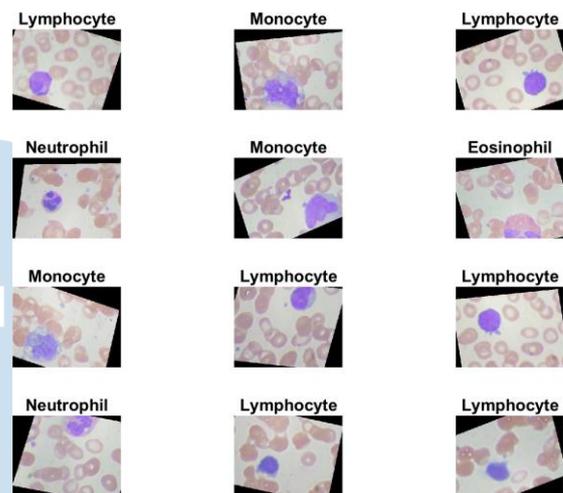
Gambar 5. Citra *Neutrophil*.

Parameter *training* yang digunakan untuk semua *network* adalah sebagai berikut: *learning rate* 1e-4, *minibatch size* 10, *max epoch* 30, *validation frequency* 3.

Arsitektur *network* diimplementasikan dengan MATLAB pada prosesor Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz dan RAM 16,0 GB dengan unit pemrosesan grafis (GPU) NVIDIA 1 GB.

IV. HASIL DAN PEMBAHASAN

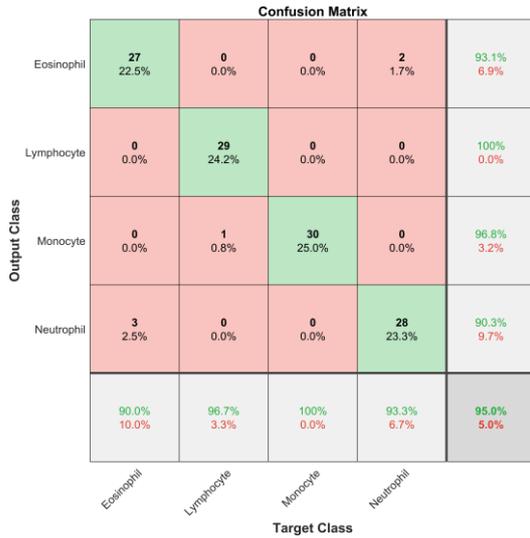
Pada eksperimen yang telah dilakukan, semua arsitektur *network* berhasil melakukan klasifikasi sel darah putih dengan menggunakan *transfer learning*. Contoh hasil klasifikasi sel darah putih ditunjukkan pada Gambar 6.



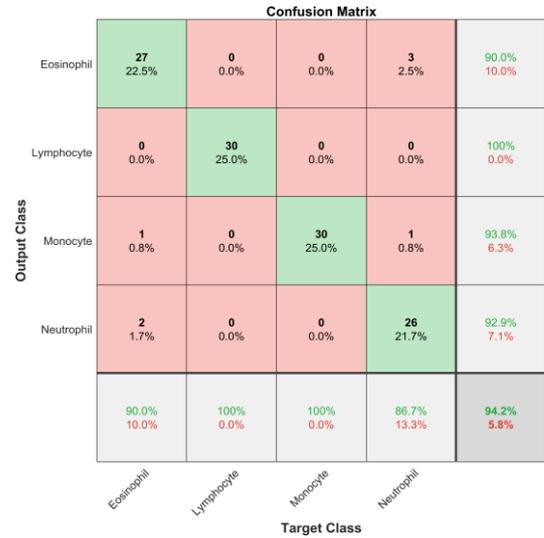
Gambar 6. Contoh hasil klasifikasi sel darah putih.

A. Perbandingan *Confusion matrix*

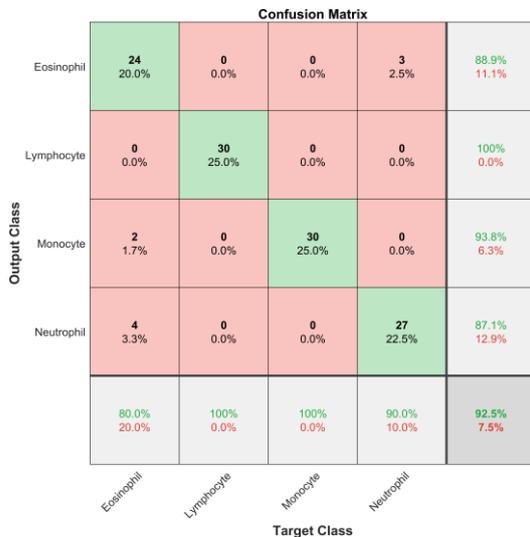
Performa klasifikasi untuk tiap arsitektur *network* disajikan dalam bentuk *Confusion matrix* yang ditunjukkan pada Gambar 7, Gambar 8, Gambar 9 dan Gambar 10. Untuk mengevaluasi hasil klasifikasi tersebut, digunakan empat *performance metrics* yaitu *Accuracy*, *Precision*, *Recall* dan *Fmeasure*. Hasil pengukuran *performance metrics* tersebut disajikan pada Tabel 2.



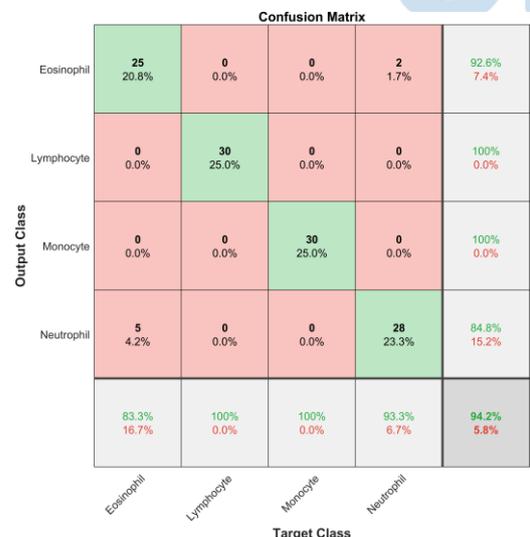
Gambar 7. Confusion matrix AlexNet.



Gambar 10. Confusion matrix VGG-16.



Gambar 8. Confusion matrix GoogleNet.



Gambar 9. Confusion matrix ResNet-50.

Nilai pada diagonal utama CM mewakili semua contoh yang diklasifikasikan dengan benar. Baris di bawah setiap CM menunjukkan tingkat *accuracy* yang dicapai untuk setiap kelas yang diprediksi dan kesalahan prediksi kelas. Gambar 7, Gambar 8, Gambar 9 dan Gambar 10 menunjukkan bahwa jumlah citra yang diklasifikasikan dengan benar pada arsitektur ResNet-50 paling tinggi dibandingkan dengan 3 arsitektur yang lain. Urutan berikutnya adalah arsitektur VGG-16, disusul oleh GoogleNet dan terakhir adalah AlexNet. Urutan yang sama juga terjadi pada perbandingan tingkat *accuracy* yang diukur pada setiap kelas pada kolom paling kanan dan paling bawah. Hal ini menunjukkan bahwa berdasarkan CM tersebut, arsitektur ResNet-50 memiliki performa yang terbaik dalam melakukan klasifikasi tersebut.

B. Evaluasi menggunakan Performance metrics

Tabel 2 menunjukkan *performance metrics* untuk 4 arsitektur *network* yang digunakan. *Performance metrics* yang digunakan adalah *Accuracy*, *Precision*, *Recall* dan *Fmeasure*.

Tabel 2. Performance metrics

	AlexNet	GoogLeNet	Resnet-50	VGG-16
Accuracy	0,9000	0,9333	0,9417	0,9167
Precision	0,8992	0,9348	0,9424	0,9287
Recall	0,9000	0,9333	0,9416	0,9166
Fmeasure	0,8996	0,9340	0,9420	0,9226

Tabel 2 menunjukkan bahwa ResNet-50 menunjukkan performa paling tinggi untuk *metric accuracy* yaitu 94,17% sedangkan GoogleNet, VGG-16, dan AlexNet masing-masing memiliki *accuracy* 93,33%, 91,67% dan 90,00%. Jika perbandingan *accuracy* ini ditinjau dari selisih *accuracy* maka ResNet-50 memiliki *accuracy* dengan selisih 0,84% terhadap GoogleNet, 2,5% terhadap VGG-16 dan 4,17% terhadap AlexNet.

Untuk 3 *metric* yang lain yaitu *Precision*, *Recall* dan *Fmeasure*, Tabel 2 juga menunjukkan urutan nilai *metric* yang sama. ResNet-50 memiliki nilai paling tinggi, disusul oleh GoogleNet, VGG-16 dan AlexNet. Dengan demikian secara keseluruhan berdasarkan CM dan *performance metrics*, urutan performa arsitektur tersebut adalah ResNet-50, GoogleNet, VGG-16, dan AlexNet.

V. SIMPULAN

Dalam penelitian ini dilakukan analisis perbandingan performa arsitektur *network* dalam mengklasifikasikan citra digital sel darah putih. Metode *transfer learning* dengan menggunakan arsitektur *network* yang telah dilatih terlebih dahulu diterapkan untuk mengklasifikasikan citra sel tersebut. Terdapat empat arsitektur *network* yang digunakan yaitu AlexNet, GoogleNet, ResNet-50 dan VGG-16. Untuk meningkatkan jumlah dataset dan mengurangi masalah *overfitting* juga digunakan augmentasi data pada metode yang diterapkan. Untuk mengevaluasi performa tiap arsitektur *network* yang diterapkan maka digunakan CM dan *performance metrics* yaitu *Accuracy*, *Precision*, *Recall* dan *Fmeasure*.

Hasil penelitian ini menunjukkan bahwa empat arsitektur *network* yang digunakan berhasil mengklasifikasikan citra sel darah putih menggunakan metode *transfer learning*. Jumlah citra yang diklasifikasikan dengan benar oleh arsitektur ResNet-50 paling tinggi dibandingkan dengan 3 arsitektur yang lain. Urutan berikutnya adalah arsitektur VGG-16, disusul oleh GoogleNet dan terakhir adalah AlexNet. Untuk *metric accuracy*, ResNet50 mencapai nilai paling tinggi yaitu 94,17% dan disusul oleh arsitektur GoogleNet, VGG-16, dan AlexNet. *Performance metrics* yang lain yaitu *Precision*, *Recall* dan *Fmeasure* menunjukkan urutan yang sama yaitu ResNet-50, GoogleNet, VGG-16 dan AlexNet. Dengan demikian ResNet-50 adalah arsitektur *network* yang menunjukkan performa paling tinggi dalam mengklasifikasikan citra sel darah putih.

Untuk arah penelitian di masa mendatang, disarankan meningkatkan *accuracy* dan nilai *performance metrics* yang lain menggunakan arsitektur *network* dengan menerapkan teknik atau algoritma optimasi.

UCAPAN TERIMA KASIH

Penulis menyampaikan terimakasih kepada Program Studi Informatika Universitas Ciputra Surabaya yang telah memberikan dukungan kepada penulis dalam melakukan penelitian ini.

DAFTAR PUSTAKA

- [1] A. T. Sahlol, P. Kollmannsberger, and A. A. Ewees, "Efficient Classification of White Blood Cell Leukemia with Improved Swarm Optimization of Deep Features," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-59215-9.
- [2] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
- [3] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Z. Med. Phys.*, vol. 29, no. 2, pp. 102–127, 2019, doi: 10.1016/j.zemedi.2018.11.002.
- [4] L. D. Nguyen, D. Lin, Z. Lin, and J. Cao, "Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation," *Proc. - IEEE Int. Symp. Circuits Syst.*, vol. 2018-May, pp. 3–7, 2018, doi: 10.1109/ISCAS.2018.8351550.
- [5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-December, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [7] C. Szegedy et al., "Going deeper with convolutions," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June-2015, pp. 1–9, 2015, doi: 10.1109/CVPR.2015.7298594.
- [8] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015, doi: 10.1007/s11263-015-0816-y.
- [9] H. C. Shin et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016, doi: 10.1109/TMI.2016.2528162.
- [10] T. Pansombut, S. Wikaisuksakul, K. Khongkraphan, and A. Phon-On, "Convolutional neural networks for recognition of lymphoblast cell images," *Comput. Intell. Neurosci.*, vol. 2019, 2019, doi: 10.1155/2019/7519603.
- [11] A. Çınar and S. A. Tuncer, "Classification of lymphocytes, monocytes, eosinophils, and neutrophils on white blood cells using hybrid Alexnet-GoogleNet-SVM," *SN Appl. Sci.*, vol. 3, no. 4, pp. 1–11, 2021, doi: 10.1007/s42452-021-04485-9.
- [12] A. Kensert, P. J. Harrison, and O. Spjuth, "Transfer learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes," *SLAS Discov.*, vol. 24, no. 4, pp. 466–475, 2019, doi: 10.1177/2472555218818756.
- [13] H. M. Ahmad, S. Ghuffar, and K. Khurshid, "Classification of Breast Cancer Histology Images Using Transfer learning," *Proc. 2019 16th Int. Bhurban Conf. Appl. Sci. Technol. IBCAST 2019*, pp. 328–332, 2019, doi: 10.1109/IBCAST.2019.8667221.
- [14] S. Hosseinzadeh Kassani and P. Hosseinzadeh Kassani, "A comparative study of deep learning architectures on melanoma detection," *Tissue Cell*, vol. 58, no. April, pp. 76–83, 2019, doi: 10.1016/j.tice.2019.04.009.
- [15] V. Maeda-Gutiérrez et al., "Comparison of convolutional neural network architectures for classification of tomato plant diseases," *Appl. Sci.*, vol. 10, no. 4, 2020, doi: 10.3390/app10041245.

Pengenalan Aktivitas Manusia Melalui Analisis Data Gerakan *Smartphone*

Eunike Endariahna S¹, Andes Suciani², Philipus Silaen³, Septian Adiwibowo⁴

^{1,2,3,4}Fakultas Ilmu Komputer, Program Studi Magister Teknologi Informasi, Universitas Indonesia, Jakarta, Indonesia

¹Fakultas Teknik dan Informatika, Program Studi Informatika, Universitas Multimedia Nusantara, Tangerang, Indonesia

¹eunike.endahriana@umn.ac.id, ²andes.suciani@ui.ac.id, ³philipus.silaen@ui.ac.id, ⁴septian.adibowo@ui.ac.id

Diterima 18 Mei 2021

Disetujui 16 Juni 2021

Abstract— Technology whose benefits are felt by humans from the development of cellular telephones is the birth of smart phones or *smartphones*. . The development of *smartphone* integration is growing very rapidly. Today's *smartphones* are not only communication tools but now offer many features and deployment of sensors. *Smartphones* are also designed to track the user's daily activities, learn and then help the user to make better decisions about what the user will take in the future. Applications that utilize the movement of a *smartphone* to analyze human activity are used by the Moves app, Fitbit Charge, Nike Fuelband, Apple Watch Health app. To perform human motion *recognition* activities, data is generated and collected from *smartphones* such as iPhones and Androids, or wearables such as the Apple Watch *smartwatch*, Nike Fuelband, and Fitbit Charge. Today the technology industry has succeeded in producing good quality devices at low prices. Sensors commonly used to collect data include an *accelerometer*, *gyroscope*, heart rate monitor, and thermometer. This study collects and manages data from the use of sensors on *smartphones*. The resulting data also combines these sensors with a magnetometer and GPS. So that we can find out the position, distance and number of points from using smart devices. The resulting data is very large, fast and real time data or what is called big data. Then these large data sets can produce complex and unstructured data. The processing requires more difficult and large computations and storage to produce useful data sets. The resulting benefits can provide problem solutions, business benefits and even new business opportunities. The resulting data can be used to seek opportunities for the benefits of human movement. The tangible benefits that can be provided are *Monitoring* city prisoners for government, community grouping, city density detection and distribution maps that can be used by business opportunities such as retail or banks.

Index Terms—*Accelerometer*, *Big Data Analytics*, *Human Movement Recognition Activity*, *Gyroscope*, *Heart Rate Monitor*, *Thermometer*.

I. PENDAHULUAN

Tujuan dari diciptakannya teknologi adalah memberikan kemudahan dan kenyamanan untuk membantu manusia. Salah satu teknologi yang

manfaatnya sangat dirasakan oleh manusia dari perkembangan telepon seluler adalah lahirnya telepon pintar atau *smartphone*. Di awal tahun 2000, belum banyak yang memiliki telepon pintar di Indonesia. Ketika itu, telepon pintar masih dianggap sebagai keperluan sekunder bahkan tersier oleh sebagian kalangan. Kini keadaannya telah berbeda, telepon pintar telah berubah menjadi kebutuhan primer yang dimiliki oleh setiap orang. Di Indonesia sendiri, negara ini berada di urutan ke enam dengan jumlah penduduknya diperkirakan sebanyak 261 juta jiwa, telah menggunakan telepon sebanyak 236 juta unit [1].

Telepon pintar kini menawarkan banyak fitur lain seperti multitasking dan penyebaran berbagai sensor, di samping sebagai alat komunikasi. Integrasi perangkat mobile ini dalam kehidupan kita sehari-hari berkembang pesat. *Smartphone* juga dirancang untuk melacak kegiatan penggunaannya setiap hari, belajar dan kemudian membantu pengguna untuk membuat keputusan yang lebih baik mengenai tindakan yang dilakukan pengguna di masa mendatang [2]. Dengan memanfaatkan gerakan dari *smartphone* untuk menganalisa aktivitas manusia yang digunakan oleh Moves app, Fitbit Charge, Nike Fuelband, Apple Watch Health app.

Aplikasi-aplikasi ini mendapatkan data dari jenis sensor terdapat pada *smartphone* seperti *accelerometer*, *gyroscope*, sensor tekanan, sensor cahaya. *Accelerometer* merupakan sensor yang sudah menjadi fitur standar atau harus ada dari *smartphone*. *Accelerometer* merupakan elemen sensor yang mengukur percepatan, sudut kemiringan, tahanan, rotasi, benturan, vibrasi dan gravitasi. Kegunaannya bisa mendukung berbagai kegiatan pengguna seperti: membaca, bermain game, olahraga dan berbagai aktivitas lainnya. Contohnya adalah Moves app dari gerakan yang dihasilkan pengguna, aplikasi ini bisa mengetahui pengguna sedang jalan, lari pelan, lari sprint, bersepeda, naik tangga, turun tangga, naik kendaraan bermotor atau sedang dalam perjalanan di pesawat. Keuntungan dari data yang diolah aplikasi ini

membantu pengguna dalam meningkatkan kesehatan seperti menghitung jumlah langkah setiap hari. Pengguna dapat melacak area yang sudah ditempuh saat berolahraga atau pun lokasi yang dia datangi setiap harinya. Berikut Gambar 1. adalah perbandingan Moves app dengan aplikasi fitness lainnya:

	Moves	Gadgets Nike+ Fuelband, Fitbit, etc.	Tracker apps Nike+ Running, Runkeeper, Runastic, etc.
Automatic activity type recognition	✓		
Calorie counter	✓	✓	✓
Routes on map	✓		✓
Daily storyline with places	✓		
No need to start and stop	✓	✓	
No need to charge and carry an extra device	✓		✓

Gambar 1. Perbandingan Moves App dengan Aplikasi Fitness Lainnya [3]

Dalam penelitian ini, kami akan meneliti lebih lanjut perkembangan data yang bisa dimanfaatkan dari gerakan *smartphone* oleh pengguna setiap harinya. Aktivitas manusia setiap harinya akan terus bertumbuh dan menjadi big data. Cara mendapatkan data dari sensor *smartphone* kemudian diolah menggunakan metode machine learning agar menghasilkan informasi yang lebih berguna untuk kedepannya.

II. METODE PENELITIAN

Untuk melakukan aktivitas pengenalan gerakan manusia, data dihasilkan dan dikumpulkan dari telepon pintar seperti iPhone dan Android, atau peralatan yang dapat dipakai seperti jam tangan pintar Apple Watch, Nike Fuelband, dan Fitbit Charge (Gambar 2) [4]. Sensor yang biasa digunakan untuk mengumpulkan data di antaranya adalah *accelerometer*, *gyroscope*, monitor detak jantung, dan termometer. Metode lain juga mengkombinasikan sensor-sensor ini dengan magnetometer dan GPS (*Global Positioning System*). [2]

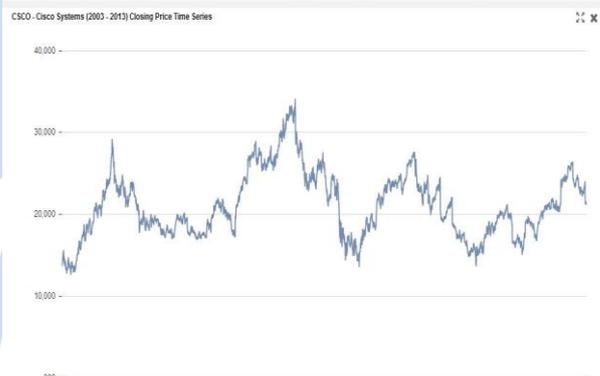


Gambar 2. Jam Tangan Pintar Fitbit Charge [5]

Data yang dihasilkan dari peralatan sensor ini berbentuk *time series* data, yaitu data yang berbentuk banyak titik data yang diurutkan berdasarkan ukuran waktu. *Time series* data ini diambil secara berurutan dengan rentang waktu atau frekuensi yang konsisten. Agar bisa mendapatkan konteks informasi, data jenis ini harus diperlakukan sebagai sebuah urutan data dan tidak bisa diperlakukan sebagai titik-titik data yang terpisah [6]. Contoh bentuk *time series* data yang umum digunakan adalah data pergerakan harga stok saham yang konteks informasinya baru bisa didapatkan apabila data ditampilkan dan diurutkan dalam suatu periode (Gambar 3).

Bab ini akan menjelaskan mengenai dua alat sensor utama yang digunakan untuk mengumpulkan data gerakan manusia, yaitu *accelerometer* dan *gyroscope* beserta contoh format data dari masing-masing peralatan tersebut. Berikutnya teknik-teknik pembersihan data (*data cleaning*) dan persiapan data (*data preparation*) yang umum dilakukan untuk kedua sensor ini akan dijabarkan. Pembahasan akan menitikberatkan ke riset utama yang dibahas di tulisan ini, yaitu penelitian dari Anguita, Ghio, Oneto, Llanas-Parra, & Reyes-Ortiz [2].

CSCO - Cisco Systems (2003 - 2013) Closing Price Time Series



Gambar 3. Time Series Data Pergerakan Harga Saham [7]

A. Accelerometer

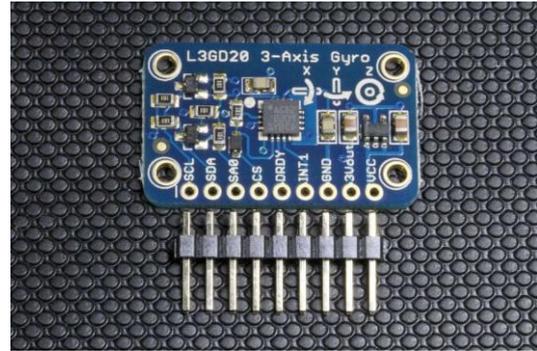
Accelerometer adalah alat untuk mengukur akselerasi atau percepatan, yaitu menghitung berapa banyak kecepatan suatu benda berubah dalam suatu rentang waktu [8]. Menggunakan prinsip tersebut, *accelerometer* bisa memberikan informasi mengenai perubahan arah dan pergerakan suatu benda.

Pada awalnya, *accelerometer* digunakan di pesawat luar angkasa (*space shuttle*) untuk mengukur perubahan kecepatan roket pesawat, mengukur *apogee* (ketika pesawat berada di titik terjauh dari bumi sehingga pengaruh akselerasi karena gravitasi berada di titik paling rendah), dan mengukur orientasi pesawat dengan cara mendeteksi sisi tekanan gravitasi bumi ketika pesawat berubah arah.

angkasa Hubble memiliki banyak *gyroscope* untuk menjaga orientasi mereka terhadap matahari [15].

Digital *gyroscope* (Gambar 7.) tidak memiliki roda bergerak, melainkan menggunakan 3 buah struktur *tuning fork* yang terbuat dari *silicon wafer*. Mereka akan bergetar bila distimulasi dengan sinyal elektrik. Ketika sumbu pergerakan benda diputar, *tuning fork* ini akan merasakan tekanan dikarenakan gaya Coriolis. Perbandingan tekanan ini proporsional dengan kecepatan rotasi benda. Ketiga buah *tuning fork* ini disusun secara ortogonal masing-masing untuk merepresentasikan sumbu X, Y, dan Z. Perputaran *gyroscope* akan direpresentasikan ke dalam data rotasi *degrees-per-second* untuk masing-masing sumbu tersebut [9] Sama seperti *accelerometer* yang pengambilan datanya menggunakan pola frekuensi, data yang dihasilkan dari *gyroscope* juga berbentuk *time series* data.

Meskipun sekilas mirip, perbedaan utama antara *accelerometer* dengan *gyroscope* adalah *accelerometer* tidak dapat mendeteksi rotasi. *Accelerometer* berguna untuk mendeteksi perpindahan, tetapi *gyroscope* bisa mendeteksi apabila rotasi terjadi tanpa perpindahan (hal yang tidak bisa dideteksi oleh *accelerometer*). Penggunaan keduanya secara bersamaan di *smartphone* dan *smartwatch* dalam frekuensi yang sama akan melengkapi satu sama lain untuk menghasilkan *dataset* yang bisa dianalisis dalam pengenalan aktivitas manusia.



Gambar 7. Digital Gyroscope [16]

C. Koleksi dan Pembersihan Data

Pada Riset Anguita, Ghio, Oneto, Llanas-Parra, & Reyes-Ortiz uji coba dilakukan dengan membuat dan mengembangkan *dataset* pengenalan aktivitas manusia dengan total sukarelawan yang berpartisipasi yaitu 30 orang dengan usia 19 hingga 48 tahun [2]. Penelitian dilakukan dengan mengumpulkan rangkaian gerakan yang terdiri dari 6 usulan Activities of Daily Living (ADL) (berdiri, duduk, berbaring, berjalan, berjalan ke atas dan berjalan turun). Setiap subjek melakukan protokol percobaan dua kali, dan setiap aktivitas setidaknya dilakukan dua kali pada setiap percobaan untuk mensimulasikan pengulangan.

Tabel 1. Protokol Percobaan [2]

No.	Static	Time (sec)	No.	Dynamic	Time (sec)
0	Start (Standing Pos)	0	7	Walk (1)	15
1	Stand (1)	15	8	Walk (2)	15
2	Sit (1)	15	9	Walk Downstairs (1)	12
3	Stand (2)	15	10	Walk Upstairs (2)	12
4	Lay Down (1)	15	11	Walk Downstairs (1)	12
5	Sit (2)	15	12	Walk Upstairs (2)	12
6	Lay Down (2)	15	13	Walk Downstairs (3)	12
			14	Walk Upstairs (3)	12
			15	Stop	0
				Total	192

Basis data yang diperoleh dipartisi menjadi data pelatihan dan data tes dalam proporsi 70% hingga 30%. Partisi diacak tetapi dipastikan bahwa tidak ada sampel yang berasal dari pengguna yang sama di kedua himpunan bagian. Data pelatihan digunakan untuk melatih *Multiclass SVM classifiers*. Percobaan diimplementasikan menggunakan *smartphone* Samsung Galaxy S II dilengkapi dengan baterai Li-Ion 1650 mAh dengan pengoperasian siaga hingga 610 jam dan sistem operasi Android *Gingerbread* versi 2.3.4.

Smartphone yang digunakan menggunakan *accelerometer*. Sinyal akselerasi dicatat pada frekuensi konstan 50Hz yang cukup cepat untuk memperoleh informasi gerak tubuh manusia. Proses pelabelan dilakukan manual dengan memilih video yang direkam dari percobaan sebagai kebenaran dasar dan membandingkannya dengan *file log* dari sinyal inersia.



Gambar 8. Pipa Proses Activity Recognition

III. HASIL DAN PEMBAHASAN

Penelitian dengan judul “Energy Efficient Smartphone-Based Activity Recognition using Fixed-point Arithmetic” [2] bertujuan untuk menerapkan aritmatika *fixed-point* untuk mengusulkan algoritma pembelajaran *Multiclass* Vector Machine yang dimodifikasi (SVM) yaitu dengan model *Multiclass* Hardware Friendly SVM (MC-HF-SVM) dengan *dataset* yang telah dibahas pada Bab 2. Berikut adalah langkah-langkah dalam model *Multiclass* HF-SVM (MC-HF-SVM):

1. Menyelesaikan masalah minimalisasi *Constrained Quadratic Programming* (CQP) dengan cara menggunakan Lagrange *multipliers* α_i . Representasi tersebut mempunyai nama lain yaitu *dual formulation* dengan rumus yaitu:

$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - r^T \alpha$$

$$0 \leq \alpha_i \leq C \forall i \in [1, \dots, l], y^T \alpha = 0 \quad (1)$$

Dimana Q adalah matriks kernel dan sebuah *semidefinite* positif simetris $l \times l$ matrix dimana $q_{ij} = y_i y_j K(x_i, x_j)$.

2. Setelah didapatkan pola yang baru maka dapat dilakukan SVM Feed-Forward Phase (FFP) dengan rumus:

$$f(x) = \sum_{i=1}^l y_i \alpha_i K(x_i, x) + b. \quad (2)$$

3. Kemudian karena adanya isu bahwa hasil yang diperoleh tidak valid dalam penggunaan *fixed-point arithmetic* karena nilai α_i masuk ke dalam kelompok bilangan real yang dibatasi antara 0 dan C, maka dilakukanlah proses normalisasi tanpa mempengaruhi tanda *output*

dari *classifier*-nya tetapi justru besarnya dan membuat akurasi dari SVMnya tetap sama. Rumus normalisasinya yaitu:

$$\beta_i = \alpha_i \frac{2^k - 1}{C} \quad (3)$$

4. Kemudian karena parameter dari *dataset* penelitian ini berupa *integer* dilakukanlah modifikasi rumus menjadi seperti berikut :

$$\begin{aligned} \min_{\beta} \frac{1}{2} \beta^T Q \beta - s^T \beta \text{ s. t. } 0 \leq \beta_i \\ \leq \frac{2^k - 1}{C} \forall i \in [1, \dots, l] \end{aligned} \quad (4)$$

5. Untuk akhirnya memiliki FFP penuh dengan hanya nilai *integer*, perlu dilakukan perubahan representasi dari kernel $K(\cdot, \cdot)$ dan vektor input x dalam hal jumlah bit (u dan v bit masing-masing). Perubahan representasi ini menghasilkan rumus:

$$0 \leq K(x_i, x) \leq 1 - 2^{-u} \forall i \in [1, \dots, l],$$

$$0 \leq x_i \leq 1 - 2^{-v} \forall i \in [1, \dots, m]. \quad (5)$$

Jika digabungkan maka berikut rumus vektor *fixed-point* FFP yang sudah dimodifikasi :

$$f(x) = \sum_{i=1}^l y_i \beta_i K(x_i, x) \quad (6)$$

6. Untuk memperluas masalah biner menjadi masalah *Multiclass*, digunakan metode OVA untuk membandingkan masing-masing kelas c terhadap kelas-kelas lain, tetapi sebelumnya sebuah prosedur untuk memungkinkan perbandingan dalam kelompok pengklasifikasi SVM diperlukan. Hasilnya, dilakukan pengukuran perkiraan probabilitas

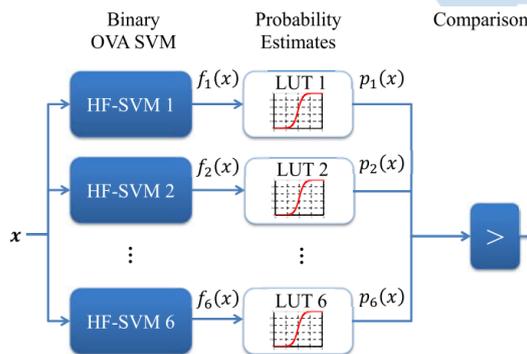
untuk setiap PC SVM (x) dan memilih kelas aktual c^* sebagai probabilitas *output* tertinggi untuk sampel uji yang diberikan

7. Kemudian *probability estimation* yang sudah didapatkan akan di implementasi menggunakan pendekatan dimana training set dan model SVM diberlakukan agar sesuai dengan nilai *output* FFP $f(x)$ dengan fungsi sigmoid dari rumus berikut :

$$p(x) = \frac{1}{1 + e^{(\Gamma f(x) + \Delta)}} \quad (7)$$

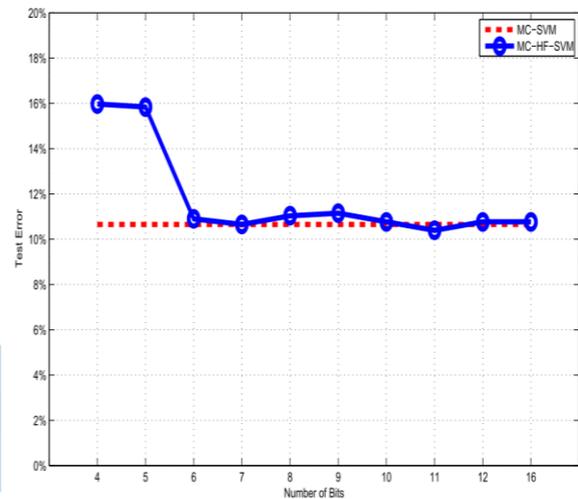
Dimana $p(x)$ adalah perkiraan probabilitas, dan Γ dan Δ adalah parameter fungsi yang dipasang dengan benar pada sampel pembelajaran yang tersedia.

8. Lalu dengan mempertimbangkan batasan aritmatik *fixed-point*, maka fungsi sigmoid yang juga bekerja dengan bilangan real, tidak dapat langsung digunakan untuk memperkirakan $p(x)$. Isu ini dapat diselesaikan dengan menggunakan Look-Up-Tables (LUTs). Pertama, sejumlah tetap bit t harus didefinisikan dan kemudian perkiraan probabilitas $p(x)$ dapat dipetakan dengan $f(x)$ tanpa perlu aritmatika *floating-point*. Proses MC-HF-SVM secara lengkap diilustrasikan dalam gambar 9.



Gambar 9. MC-HF-SVM Menggunakan LUT Untuk Probability Estimates [2]

Kemudian dari langkah-langkah tersebut menunjukkan bahwa $t = 8$ adalah nilai yang cocok untuk aplikasi ini dan hanya dibutuhkan LUT dengan 256 elemen. Lalu dilakukan evaluasi dengan membandingkan hasil *error rates* dan hasil *classification* pada test data dari MC-HF-SVM dan MC-SVM. Berikut perbandingan hasil *error rates* yang didapatkan dari kedua pendekatan tersebut dengan *range bits* dari k antara 4 sampai 16 bits:



Gambar 17. Perbandingan Antara Tingkat Kesalahan yang Diperoleh Dengan MC-SVM (Garis Putus-Putus Merah) dan MC-HF-SVM (Garis Biru) Karena k Bervariasi [2]

Dari tes yang dilakukan pada bits ke 6 sampai 16 test *error* yang ada cukup stabil (hampir 1% variasi). Ketika test dilakukan pada bits dibawah 6 maka test *error* naik secara signifikan sebanyak kira-kira dua kali lipat. Selain itu, juga terlihat dari grafik bahwa beberapa test *error* dengan pendekatan *fixed-point* lebih kecil dari yang ditemukan dengan pendekatan MC-SVM Kemudian perbandingan hasil *classification* dari kedua pendekatan tersebut dengan menggunakan $k=8$ bits dapat dilukiskan melalui tabel berikut:

Tabel 2. Confusion matrix Dari Hasil Classification Pada Test Data Menggunakan MC-SVM dan MC-HF-SVM

Method	MC-SVM							MC-HF-SVM $k = 8$ bits						
	Walking	Upstairs	Downstairs	Standing	Sitting	Laying	Recall %	Walking	Upstairs	Downstairs	Standing	Sitting	Laying	Recall %
Walking	109	0	5	0	0	0	95.6	109	2	3	0	0	0	95.6
Upstairs	1	95	40	0	0	0	69.8	1	98	37	0	0	0	72.1
Downstairs	15	9	119	0	0	0	83.2	15	14	114	0	0	0	79.7
Standing	0	5	0	132	5	0	93.0	0	5	0	131	6	0	92.2
Sitting	0	0	0	4	108	0	96.4	0	1	0	3	108	0	96.4
Laying	0	0	0	0	0	142	100	0	0	0	0	0	142	100
Precision %	87.2	87.2	72.6	97.1	95.6	100	89.3	87.2	81.7	74.0	97.8	94.7	100	89.0

(Sumber: Anguita, Ghio, Oneto, Llanas-Parra, & Reyes-Ortiz, 2013)

Tabel 2. merupakan *Confusion matrix* dari hasil *classification* pada test data menggunakan *traditional floating point* MC-SVM (kiri) dan MC-HF-SVM dengan $k = 8$ bits (kanan). Baris mewakili kelas yang sebenarnya dan kolom mewakili kelas yang terprediksi. Angka yang di bold secara diagonal menunjukkan jumlah sampel tes yang dengan benar diklasifikasikan. Pada tabel 2. dilakukan perbandingan yang mengukur keseluruhan dari *accuracy*, *recall* dan *precision* dari kedua pendekatan tersebut.

Variasi kecil diperhatikan dalam *recognition accuracy* dari aktivitas dinamis dalam dua pendekatan tersebut, seperti dalam kegiatan *walking downstairs* dan *walking upstairs*, yang juga menampilkan beberapa kesalahan *classification* terutama untuk kesamaan gerakan mereka. Aktivitas statis di sisi lain berkinerja lebih baik, seperti *laying*, *classification error* yang didapat sebesar 0%. Selain itu, ada kesalahan klasifikasi kecil yang ditemukan antara *standing* dan *sitting*, yang dikaitkan dengan lokasi fisik telepon seluler pintar yang dipasang di pinggang dan kesulitan untuk membedakan di antara kedua aktivitas tersebut.

Hasil dari penelitian ini yaitu model *Multiclass SVM* standar yang dapat diubah dengan representasi *fixed-point* yang lebih efisien. Keuntungan dari penerapan model yang diusulkan dibandingkan dengan pendekatan tradisional SVM yang lain yaitu waktu proses yang lebih cepat, dan penggunaan sumber daya sistem yang lebih sedikit dimana hasilnya memberikan penghematan dalam konsumsi energi baterai perangkat dan juga kemungkinan untuk menggunakan pendekatan ini pada perangkat dengan harga murah (misalnya perangkat keras *fixed-point*).

IV. KESIMPULAN

Sistem pengenalan aktivitas manusia dikembangkan sebagai bagian dari kerangka kerja untuk memungkinkan pemantauan terus menerus perilaku manusia di bidang kehidupan dibantu lingkungan, deteksi cedera olahraga, perawatan lansia, rehabilitasi, hiburan dan pengawasan di lingkungan rumah pintar. Proses pembelajaran mesin yang relevan adalah bagian yang paling menantang dari pengenalan aktivitas manusia yang dapat diaktifkan dan berdasarkan sensor. Metode yang digunakan juga berbagai macam dalam mengolah data.

Berdasarkan metode cara pembelajaran mengelola data aktivitas manusia, jika ingin dikembangkan teknologi ini bisa diterapkan pada:

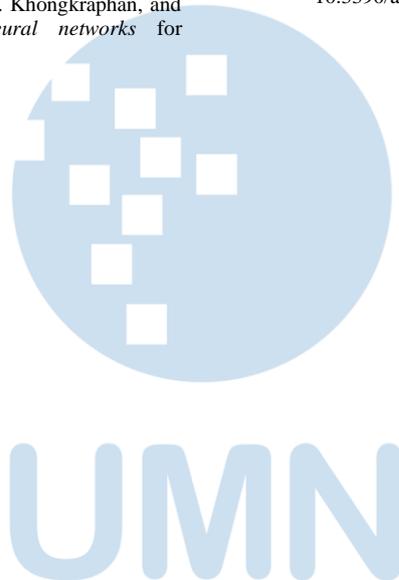
1. *Monitoring* tahanan kota, tahanan kota yang ditahan tidak boleh meninggalkan satu kota yang menjadi tempat penahanannya. Pengawasannya, tahanan wajib melapor dua kali dalam seminggu. Dengan memberikan alat pendeteksi aktivitas

- manusia ini, membantu kepolisian dalam *monitoring* keberadaan tahanan setiap waktu.
2. *Monitoring* Orang Tanpa Gejala (OTG) penderita COVID-19 yang melakukan isolasi mandiri. Dengan pengawasan ini, OTG dapat dipantau posisinya. Jika keluar dari area yang di ijin maka sensor pada *smartwatch* memberikan notifikasi peringatan awal atau memberitahukan tim medis yang bertanggung jawab dalam *monitoring* pasien yang melakukan isolasi mandiri.
3. Pendeteksian komunitas, adalah cara untuk memahami struktur jaringan yang kompleks, untuk akhirnya mengekstrak informasi yang berguna yang bisa didapatkan. Aplikasi beragam mulai dari perawatan, kesehatan, geografi regional, interaksi manusia, mobilitas hingga ekonomi [17][18].
4. Kepadatan kota dan peta distribusi, informasi yang diperoleh mengetahui titik dimana orang banyak berkumpul. Dari aktivitas yang mereka lakukan, kita dapat informasi bukan hanya gender, tetapi rentang usia, dan kapan tanggal aktivitas ini dilakukan. Pola yang didapatkan ini dapat digunakan untuk dijual kepada pelaku bisnis seperti Warunk Upnormal, Indomaret, Alfamart, pom bensin dan lain lain dalam membuka cabang baru.
5. Di tengah kondisi pandemi ini, kita bisa mendeteksi kerumunan padat suatu daerah sehingga bisa dicapai langkah awal pembubaran masa. Selain itu pendeteksian suhu, detak jantung, tekanan darah dan alat vital lainnya di tubuh manusia dapat diketahui dengan penggunaan sensor pada *smartwatch* dan terintegrasi pada *smartphone*. Mendeteksi alat vital pasien dengan menggunakan sensor dapat mengurangi kontak fisik antara pasien dengan dokter dan tim medis. Data alat vital pasien menjadi syarat awal ketika ingin melakukan pemeriksaan lebih lanjut. Pendeteksian ini juga bisa membantu mengurangi tingkat penyebaran virus COVID 19.

DAFTAR PUSTAKA

- [1] A. T. Sahlol, P. Kollmannsberger, and A. A. Ewees, "Efficient Classification of White Blood Cell Leukemia with Improved Swarm Optimization of Deep Features," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, 2020, doi: 10.1038/s41598-020-59215-9.
- [2] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.
- [3] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Z. Med. Phys.*, vol. 29, no. 2, pp. 102–127, 2019, doi: 10.1016/j.zemedi.2018.11.002.
- [4] L. D. Nguyen, D. Lin, Z. Lin, and J. Cao, "Deep CNNs for microscopic image classification by exploiting transfer

- learning and feature concatenation,” *Proc. - IEEE Int. Symp. Circuits Syst.*, vol. 2018-May, pp. 3–7, 2018, doi: 10.1109/ISCAS.2018.8351550.
- [5] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual learning for image recognition,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-December, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [7] C. Szegedy *et al.*, “Going deeper with convolutions,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June-2015, pp. 1–9, 2015, doi: 10.1109/CVPR.2015.7298594.
- [8] O. Russakovsky *et al.*, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015, doi: 10.1007/s11263-015-0816-y.
- [9] H. C. Shin *et al.*, “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer learning,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016, doi: 10.1109/TMI.2016.2528162.
- [10] T. Pansombut, S. Wikaisuksakul, K. Khongkrapan, and A. Phon-On, “Convolutional neural networks for recognition of lymphoblast cell images,” *Comput. Intell. Neurosci.*, vol. 2019, 2019, doi: 10.1155/2019/7519603.
- [11] A. Çınar and S. A. Tuncer, “Classification of lymphocytes, monocytes, eosinophils, and neutrophils on white blood cells using hybrid Alexnet-GoogleNet-SVM,” *SN Appl. Sci.*, vol. 3, no. 4, pp. 1–11, 2021, doi: 10.1007/s42452-021-04485-9.
- [12] A. Kensert, P. J. Harrison, and O. Spjuth, “Transfer learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes,” *SLAS Discov.*, vol. 24, no. 4, pp. 466–475, 2019, doi: 10.1177/2472555218818756.
- [13] H. M. Ahmad, S. Ghuffar, and K. Khurshid, “Classification of Breast Cancer Histology Images Using Transfer learning,” *Proc. 2019 16th Int. Bhurban Conf. Appl. Sci. Technol. IBCAST 2019*, pp. 328–332, 2019, doi: 10.1109/IBCAST.2019.8667221.
- [14] S. Hosseinzadeh Kassani and P. Hosseinzadeh Kassani, “A comparative study of deep learning architectures on melanoma detection,” *Tissue Cell*, vol. 58, no. April, pp. 76–83, 2019, doi: 10.1016/j.tice.2019.04.009.
- [15] V. Maeda-Gutiérrez *et al.*, “Comparison of convolutional neural network architectures for classification of tomato plant diseases,” *Appl. Sci.*, vol. 10, no. 4, 2020, doi: 10.3390/app10041245.



Prediksi Akurasi Kemenangan Pada Permainan Poker Menggunakan Algoritma C5.0 Dan WIPSO

M. Fariz Januarsyah¹, Dr. Ermatita, M.Kom.²

^{1,2}Program Studi Magister Ilmu Komputer, Fakultas Ilmu Komputer
Universitas Sriwijaya, Palembang, Indonesia

¹m.fariz.januarsyah@gmail.com, ²ermatita@unsri.ac.id

Diterima 24 Mei 2021

Disetujui 07 Juni 2021

Abstract—In the era of information technology, a lot of data can be taken from human activities based on computer systems. But the system is not only found on computers, but in all areas of human life, be it in terms of health, security, even in games where the data set from these activities becomes a database that can be used to find a new knowledge. This study aims to predict the accuracy of poker games using the Weight Improved Particle Swarm Optimization (WIPSO) algorithm for attribute selection which then uses the C5.0 algorithm to predict accuracy. Before being processed, the dataset will be changed from 11 attributes to 6 attributes. The results of this study indicate that the accuracy of the poker card will increase, when using the C5.0 algorithm the accuracy obtained is 49.952% while the accuracy obtained by the C5.0 + WIPSO algorithm is 51.2%.

Index Terms—Poker, Data mining, Algoritma C5.0, WIPSO.

I. PENDAHULUAN

Dalam era teknologi informasi banyak data yang dapat diambil dari kegiatan dilakukan manusia yang berhubungan dengan sistem komputer. Tetapi sistem itu bukan hanya terdapat pada komputer saja melainkan di seluruh bidang kehidupan manusia, baik itu dari segi kesehatan, keamanan bahkan di dalam permainan dengan kumpulan data dari kegiatan itu dapat menjadi *database* yang besar. *Database* ini hendaknya dapat digunakan untuk menemukan pengetahuan yang baru dalam mengambil keputusan dengan peran sebagai estimasi dan prediksi.

Permainan adalah kegiatan yang tidak dapat dilepaskan dari kehidupan manusia. Baik itu permainan fisik seperti sepak bola, atau permainan yang menggunakan strategi seperti catur. Dalam hal strategi, permainan poker merupakan salah satu permainan kartu yang mengandalkan hal tersebut [1], [2] persaingan ketat terjadi di dalam permainan poker membuat para pemainnya untuk selalu memikirkan strategi-strategi untuk menjamin akurasi kemenangan mereka [2]. Di dalam permainan pasti ada menang dan

kalah, tetapi semua orang pasti ingin memenangkan permainan mereka. Untuk mencapai hal itu didapatkanlah sebuah solusi untuk mencapai kemenangan tersebut dalam bentuk sebuah prediksi. Prediksi kemenangan telah diterapkan diberbagai jenis permainan, seperti mencari prediksi kemenangan menggunakan metode *symbolic model checking* yang di optimasi dengan metode *binary decision diagram* untuk mencari strategi kemenangan dengan hasil gabungan dari kedua metode ini berhasil memprediksi strategi memenangkan permainan catur [3], dan penelitian yang menggunakan *artificial intelligence* (AI) dengan algoritma *Union-Find Set* dan *Deep-First-Search* dalam mencari keputusan untuk memenangkan permainan berbasis *Hex Game System* (permainan papan dengan grid berbentuk heksagon) dimana penelitian ini mencari langkah terbaik berdasarkan tiga skenario yang telah ditetapkan [4].

Untuk permainan poker, beberapa penelitian telah dilakukan diantaranya adalah dengan menggunakan algoritma naive bayes yang telah dimodifikasi dan analisis statistik, didapatkan hasil prediksi kemenangan berdasarkan kombinasi kartu yang ada dengan tingkat akurasi mencapai 92,13% [5]. Hal prediksi juga dapat digunakan untuk memprediksi langkah yang akan diambil pemain dengan menggunakan *Sequential Model* dan *TensorFlow*, hasil dari penelitian ini adalah sebuah model *neural network model training* dengan tiga layer aktivasi berupa *relu*, *relu6* dan *selu*, dengan menggunakan algoritma yang dapat mengubah parameter secara otomatis sehingga akurasi prediksi yang di dapat mencapai 89,09% hingga 99,31% [6]. Penelitian selanjutnya menggunakan algoritma Bayesian Networks dan *Monte-Carlo Tree Search* untuk membuat simulasi langkah pemain poker berdasarkan model probabilistik lawan [7]. Untuk membuat sebuah prediksi, salah satu algoritma yang sering digunakan adalah Algoritma C5.0.

Algoritma C5.0 ini sudah digunakan pada penelitian-penelitian sebelumnya untuk mencari sebuah prediksi, dari bidang medis, algoritma C5.0 digunakan untuk memprediksi cacat setelah kelahiran berdasarkan data penyakit turunan, hipertensi diabetes dan lain-lain, dengan menggunakan C4.5 sebagai pembandingan, akurasi yang di dapat oleh algoritma C5.0 mencapai 94,15% dan C4.5 mencapai 90,87% [8]. Penggunaan algoritma C5.0 juga dapat digabungkan dengan algoritma-algoritma yang lain sehingga membentuk algoritma hybrid, salah satunya adalah optimasi algoritma C5.0 menggunakan teori bayesian posterior, hasil dari penelitian ini adalah *post pruning decision tree* yang mengurangi resiko *overfitting*, sebelum di optimasi algoritma C5.0 menghasilkan akurasi 85,64% dan setelah algoritma itu di optimasi menghasilkan akurasi 91,25% [9].

Ada banyak sekali metode yang dapat digunakan untuk mengoptimasi hasil dari sebuah algoritma, salah satu metode itu adalah *Particle Swarm Optimization* (PSO). Algoritma PSO ini diperkenalkan oleh Kennedy, Eberhart, dan Shi yang pertama sekali digunakan untuk mensimulasi tingkah laku sosial dari sekumpulan burung atau ikan [10], [11]. Algoritma PSO juga dapat digunakan dalam hal memprediksi dengan menggabungkan PSO dan *Neural Network* untuk memprediksi kekuatan semen pada stopes, dengan algoritma PSO bertindak sebagai *architecture-tuning* untuk *Neural Network* [12]. Sedangkan pada bidang permainan agloritma PSO juga digunakan untuk meningkatkan akurasi dari *support vector machine* (SVM) untuk menentukan pemenang dipertandingan DOTA 2 dengan PSO bertindak sebagai penentu parameter optimum untuk SVM [13].

Sejak pertama kali dikenalkan algoritma PSO ini telah menghasilkan varian-varian algoritma PSO. Salah satu varian algoritma ini adalah *Weight Improved Particle Swarm Optimization* (WIPSO) sebagai pengembangan dari algoritma PSO standar. Algoritma WIPSO digunakan pada penelitian untuk reduksi harmonik *H-Bridge* secara selektif pada multilevel inverter berdasarkan sumber yang setara dengan tujuan mengurangi *total harmonic distortion* (THD) [14], jika dibandingkan dengan algoritma PSO standar, WIPSO memiliki kelebihan dapat menemukan jawaban dengan iterasi yang lebih sedikit dengan kecepatan yang lebih tinggi.

Berdasarkan uraian di atas peneliti akan mencoba untuk menggunakan algoritma C5.0 dan algoritma WIPSO yang merupakan pengembangan dari algoritma standar PSO untuk memprediksi akurasi pada permainan poker dengan tujuan, dari penggunaan WIPSO dapat meningkatkan akurasi algoritma C5.0 dalam mencari akurasi pada permainan poker. Hasil dari penelitian ini akan di buat perbandingan akurasi antara algoritma C5.0 setelah di optimasi WIPSO dengan algoritma C5.0 sebelum optimasi WIPSO.

II. METODE PENELITIAN

A. Data Mining

Data mining adalah proses untuk menemukan pola yang bermanfaat bagi pengguna dari kumpulan data yang sangat banyak yang kemudian di olah dengan baik dengan teknik pembelajaran mesin, pengenalan pola, statistik dan visualisasi [15].

Kegiatan yang dilakukan manusia dipenuhi dengan data-data dibanyak bidang. Tetapi data yang banyak ini menghabiskan biaya yang banyak untuk di kumpul dan di susun dan data ini banyak yang tertimbun di dalam repositori. Karena itu penggunaan data mining dilakukan untuk mengubah data yang tertimbun itu menjadi sebuah pengetahuan baru yang dapat membantu kegiatan yang akan datang.

B. Algoritma C5.0

Algoritma C5.0 merupakan produk komersial hasil pengembangan dari algoritma C4.5 oleh Rulequest Research Inc [16]. C5.0 merupakan algoritma klasifikasi di dalam pembelajaran mesin dan data mining. Algoritma ini memiliki keunggulan dibandingkan dengan algoritma C4.5 baik itu dalam segi kemampuan ataupun efisiensi.

Di dalam algoritma C5.0 terdapat beberapa persamaan yang digunakan diantaranya adalah untuk mencari nilai entropi, nilai subset dan gain, untuk mencari nilai entropy di dalam algoritma C5.0 digunakan persamaan sebagai berikut. [17]

$$Entropi (s) = - \sum_{i=1}^m p_i \log_2 (p_i) \quad (1)$$

Dimana

s = himpunan (dataset) kasus

m = banyaknya partisi

p_i = probabilitas dari total kelas tertentu dibagi total kasus

Lalu nilai subset di dalam algoritma C5.0 dapat dicari dengan menggunakan persamaan sebagai berikut [17]

$$subset (a) = \sum_{i=1}^m \frac{|s_i|}{|s|} Entropi (s_i) \quad (2)$$

Dimana

s = data sampel training

a = atribut data

$|s_i|$ = jumlah sampel untuk nilai i

$|s|$ = jumlah seluruh sampel data

Entropi(s) = entropi untuk sampel yang memiliki nilai i

C. Weight Improved Particle Swarm Optimization (WIPSO) Mempertahankan Keutuhan Format

Weight Improved Particle Swarm Optimization (WIPSO) merupakan salah satu varian dari sekian banyaknya algoritma PSO, algoritma WIPSO ini meningkatkan efisiensi dari algoritma PSO standar dengan meningkatkan perhitungan faktor berat inersia, kognitif dan komponen sosial (C_1 , C_2), karena itu perhitungan di dalam algoritma ini telah berubah,

maka rumus untuk mencari kecepatan yang baru adalah [14]

$$V_{i+1} = W_{new} \times V_i + C_1 \times r_1 \times (P_{best_i} - S_i) + C_2 \times r_2 \times (g_{best} - S_i) \quad (3)$$

Nilai W_{new} dapat dicari dengan rumus sebagai berikut [14]:

$$W_{new} = W_{min} + w \times r_1 \quad (4)$$

Sedangkan nilai w dapat dicari dengan rumus [14]:

$$w = W_{max} - [(W_{max} - W_{min}) \times ITE] / (ITE_{max}) \quad (5)$$

Untuk nilai W_{max} dan W_{min} umumnya digunakan nilai $W_{max} = 0,9$ dan nilai $W_{min} = 0,4$ [18]. Selanjutnya untuk mencari nilai C_1 dan C_2 dapat digunakan persamaan sebagai berikut [14]:

$$C_1 = C_{1max} - (C_{1max} - C_{1min}) \times (ITE / ITE_{max}) \quad (6)$$

$$C_2 = C_{2max} - (C_{2max} - C_{2min}) \times (ITE / ITE_{max}) \quad (7)$$

Untuk nilai C_{1max} dan C_{2max} menggunakan nilai 2 dengan nilai C_{1min} dan C_{2min} menggunakan nilai 0 [18]. Persamaan di atas dapat dijelaskan dengan dengan keterangan sebagai berikut:

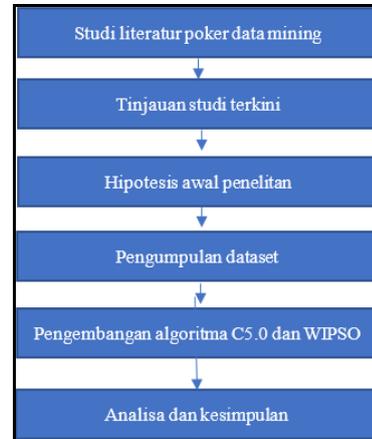
V_i	= Kecepatan partikel
S_i	= Posisi partikel saat ini
w	= Berat inersia
W_{min}	= Nilai inersia minimum
W_{max}	= Nilai inersia maksimum
C_1	= Koefisien akselerasi kognitif
C_{1min}	= Nilai awal akselerasi kognitif
C_{1max}	= Nilai final akselerasi kognitif
C_2	= Koefisien akselerasi sosial
C_{2min}	= Nilai awal akselerasi sosial
C_{2max}	= Nilai awal akselerasi sosial
P_{best_i}	= Posisi terbaik partikel
g_{best}	= Posisi terbaik global diantara partikel
r_1 dan r_2	= nilai acak antara 0 sampai 1
ITE	= Nilai pengulangan saat ini
ITE _{max}	= maksimum nilai pengulangan

Pengujian akan dilakukan dengan menggunakan algoritma WIPSO untuk seleksi atribut terlebih dahulu yang kemudian dilanjutkan dengan algoritma C5.0. Hasil dari pemrosesan ini berupa akurasi, error dan kecepatan eksekusi yang kemudian akan digunakan sebagai data pembanding. dalam pengolahan data ini akan ditentukan beberapa atribut yang akan digunakan sebagai parameter dalam pengklasifikasian data sampel. Beberapa parameter ditentukan diawal, yaitu :

- Jumlah partikel ditentukan sebanyak 50 hingga 300.
- Nilai C_1 dan C_2 antara 0,5 hingga 2.
- Iterasi ditentukan antara 100 hingga 1000 kali.
- Metode validasi adalah K-fold cross validation dengan nilai 10.

D. Tahapan Penelitian

Penelitian ini terdiri dari 6 (enam) tahap, diantaranya adalah studi literatur, tinjauan studi, hipotesis, pengumpulan data, pengembangan algoritma, analisa dan kesimpulan. Tahapan penelitian yang digunakan dan dapat dilihat pada Gambar 1



Gambar 1 Tahapan Penelitian

- Tahapan pertama pada penelitian ini adalah mencari studi literatur dan data-data yang terkait dengan permainan poker.
- Tahapan kedua adalah mencari studi terkini yang berkaitan dengan prediksi kemenangan yang ada pada bidang permainan.
- Tahapan ketiga adalah menentukan hipotesis awal dari penelitian ini, yaitu dengan menggunakan algoritma WIPSO dapat meningkatkan akurasi dari algoritma C5.0 dalam memprediksi.
- Tahapan keempat adalah mengumpulkan dataset poker dari situs <https://archive.ics.uci.edu/ml/index.php> dengan dataset poker yang digunakan adalah *poker-hand-training-true.data* yang terdiri dari 25010 data dan *poker-hand-testing.data* yang terdiri dari 1000000 data.
- Tahapan kelima adalah mengembangkan algoritma C5.0 dan algoritma WIPSO
- Tahapan keenam adalah hasil dan kesimpulan dari penelitian.

III. HASIL DAN PEMBAHASAN

A. Normalisasi Data

Sebelum dataset dapat digunakan akan dilakukan proses normalisasi data dengan menggabungkan dua atribut yang ada di dalam dataset. Dataset yang digunakan adalah *poker-hand-training-true.data* yang terdiri dari 25.010 data dan *poker-hand-testing.data* dengan 1.000.000 data. Dataset tersebut memiliki 11 atribut dimana atribut 1 hingga 10 berupa

tipe dan nilai kartu dan atribut ke-11 sebagai atribut kelas. Normalisasi yang dilakukan adalah menggabungkan kedua atribut diantara sepuluh atribut menjadi satu atribut untuk memudahkan proses sehingga bentuk dataset tersebut menjadi 6 (enam) atribut, dimana atribut 1 hingga 5 adalah atribut kartu dan atribut ke-6 adalah atribut kelas.

Pada tabel 1 adalah sampel tabel yang ada pada dataset poker dengan 11 atribut, adapun keterangan dari nilai pada atribut 'tipe' yang ada pada dataset tersebut adalah, nilai 1 pada atribut tipe menandakan tipe kartu 'Hati' (♥), nilai 2 menandakan tipe kartu 'Sekop' (♠), nilai 3 menandakan tipe kartu 'Wajik' (♦), dan nilai 4 menandakan kartu 'Keriting' (♣). Sedangkan untuk atribut 'nilai' untuk nilai 1 adalah kartu AS, sedangkan nilai 11 adalah kartu 'Jack', nilai 12 adalah kartu 'Queen', dan nilai 13 adalah kartu 'King'.

Sedangkan keterangan untuk nilai pada atribut poker hand dapat dapat disimpulkan menjadi, nilai 0 adalah pola kartu 'Nothing' atau 'High Hand', nilai 1 adalah pola kartu 'One Pair', nilai 2 adalah 'Two Pairs', nilai 3 adalah 'Three of a kind', nilai 4 adalah 'Straight', nilai 5 adalah 'Flush', nilai 6 adalah 'Full House', nilai 7 adalah 'Four of a kind', nilai 8 adalah 'Straight Flush', nilai 9 adalah 'Royal Flush' dengan nilai yang memiliki nilai besar dapat mengalahkan nilai yang ada dibawahnya, dalam hal ini nilai 0 adalah nilai terkecil yang ada pada atribut 'Poker Hand'.

Tabel 1 Data sebelum Normalisasi

T1	N1	T2	N2	T3	N3	T4	N4	T5	N5	H
1	10	1	11	1	13	1	12	1	1	9
2	11	2	13	2	10	2	12	2	1	9
3	12	3	11	3	13	3	10	3	1	9
4	10	4	11	4	1	4	13	4	12	9
4	1	4	13	4	12	4	11	4	10	9
1	2	1	4	1	5	1	3	1	6	8
1	9	1	12	1	10	1	11	1	13	8
2	1	2	2	2	3	2	4	2	5	8
3	5	3	6	3	9	3	7	3	8	8
4	1	4	4	4	2	4	3	4	5	8
1	1	2	1	3	9	1	5	2	3	1
2	6	2	1	4	13	2	4	4	9	0
1	10	4	6	1	2	1	1	3	8	0
2	13	2	1	4	4	1	5	2	11	0
3	8	4	12	3	9	4	2	3	2	1

Setelah dataset tersebut dilakukan proses normalisasi data, hasil dari proses tersebut dapat dilihat pada tabel 2. dimana dari 11 atribut data menjadi 6 atribut.

Tabel 2 Data setelah Normalisasi

Kartu 1	Kartu 2	Kartu 3	Kartu 4	Kartu 5	Poker Hand
110	111	113	112	101	9
211	213	210	212	201	9
312	311	313	310	301	9
410	411	401	413	412	9
401	413	412	411	410	9
102	104	105	103	106	8
109	112	110	111	113	8
201	202	203	204	205	8
305	306	309	307	308	8
401	404	402	403	405	8
101	201	309	105	203	1
206	201	413	204	409	0
110	406	102	101	308	0
213	201	404	105	211	0
308	412	309	402	302	1

B. Penerapan Algoritma C5.0

Setelah dataset dinormalisasi, langkah selanjutnya adalah melakukan proses eksperimen dengan menggunakan algoritma C5.0 sebelum di optimasi dengan WIPSO yang dapat dilihat pada gambar 2.

```

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      12493      49.952 %
Incorrectly Classified Instances    12517      50.048 %
Kappa statistic                     0
Mean absolute error                 0.1136
Root mean squared error             0.2384
Relative absolute error             99.9741 %
Root relative squared error         100 %
Total Number of Instances          25010
    
```

Gambar 2. Hasil Algoritma C5.0

Dari gambar tersebut dapat dilihat bahwa tingkat akurasi yang dicapai algoritma C5.0 berkisar pada 49,952% dengan menggunakan cross-validation sebagai metode validasi dari 25010 dataset yang digunakan.

C. Penerapan Algoritma C5.0 + WIPSO

Tahapan selanjutnya adalah menggunakan algoritma WIPSO sebagai seleksi atribut yang selanjutnya diproses dengan algoritma C5.0 pada data sampel. Hasil dari eksperimen tersebut dapat dilihat pada gambar 3 sebagai berikut.

```

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      52         52 %
Incorrectly Classified Instances    48         48 %
Kappa statistic                     0
Mean absolute error                 0.2238
Root mean squared error             0.3348
Relative absolute error             97.7884 %
Root relative squared error         99.9482 %
Total Number of Instances          100
    
```

Gambar 3 Hasil Algoritma C5.0 + WIPSO

Dari gambar tersebut dengan menggunakan metode validasi yang sama yaitu cross-validation dapat dilihat terjadinya peningkatan akurasi dari 49,952% menjadi 52%. Tetapi karena hasil ini masih menggunakan data sampel, langkah selanjutnya adalah melatih hasil dari algoritma ini dengan dataset poker yang dapat dilihat pada gambar 4.

```

=== Evaluation on test set ===

Time taken to test model on supplied test set: 0.07 seconds

=== Summary ===

Correctly Classified Instances      12498      51.226 %
Incorrectly Classified Instances    11895      48.774 %
Kappa statistic                    0
Mean absolute error                0.2212
Root mean squared error            0.3308
Relative absolute error            97.9165 %
Root relative squared error        99.8991 %
Total Number of Instances         24398
Ignored Class Unknown Instances    622
  
```

Gambar 4 Hasil Latih C5.0 + WIPSO

Dari gambar 4 akurasi yang didapatkan tetap meningkat dengan nilai 51,2%, hal ini membuktikan bahwa dengan menggabungkan Algoritma WIPSO sebagai algoritma optimasi dapat meningkatkan akurasi dari algoritma C5.0.

Penerapan metode ini dilakukan pada 100 hingga 1000 data sampel. Akurasi yang didapatkan dari eksperimen ini dapat dilihat pada tabel 3 sebagai berikut.

Tabel 3 Hasil eksperimen

Jumlah Sampel Data	Akurasi C5.0+WIPSO	Hasil Latih Data	Selisih Akurasi
100	52,0000%	51,2260%	1,2740%
200	49,0000%	50,3425%	0,3905%
300	52,0000%	50,1646%	0,2126%
400	50,5000%	50,2696%	0,3176%
500	53,4000%	50,1546%	0,2026%
600	50,1667%	50,1606%	0,2086%
700	49,5714%	49,9740%	0,0220%
800	49,3750%	49,9840%	0,0320%
900	53,3333%	49,9720%	0,0200%
1000	49,0000%	49,9840%	0,0320%

Seperti yang dilihat pada tabel 3 akurasi yang didapatkan untuk semua data sampel tetap meningkat, untuk perhitungan akurasi C5.0 + WIPSO, akurasi dari sampel data dengan jumlah data sebesar 500 dan 900 memiliki akurasi tertinggi yang disusul oleh 100 dan 300. Tetapi setelah dataset tersebut dilatih terjadi penurunan akurasi yang signifikan pada sampel 300, 500 dan 900, sedangkan akurasi dataset dengan 100 sampel tidak mengalami penurunan yang signifikan, sampel data dengan 100 data merupakan yang terbaik berdasarkan selisih dari hasil latih data dengan akurasi algoritma C5.0 dengan nilai 1,2740%.

IV. SIMPULAN

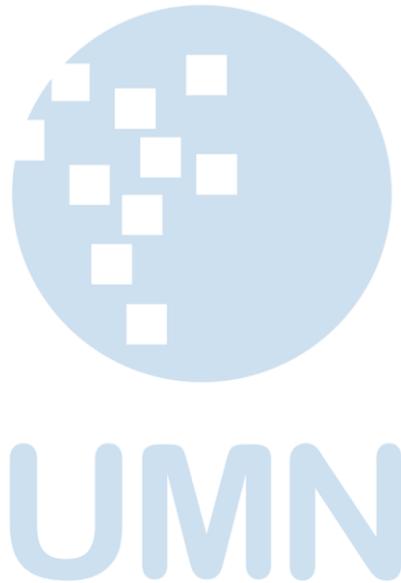
Hasil dari penelitian ini menghasilkan beberapa kesimpulan, yaitu :

1. Algoritma C5.0 dapat digunakan untuk melakukan prediksi akurasi poker dengan algoritma *Weight Improved Particle Swarm Optimization* (WIPSO) sebagai optimasi untuk meningkatkan akurasi algoritma C5.0 yang terbukti dengan adanya peningkatan sebesar 1,2740% dibandingkan dengan penggunaan algoritma C5.0 saja. Penggunaan algoritma C5.0 dari dataset tersebut mendapatkan hasil akurasi sebesar 49,952% sedangkan algoritma C5.0 + WIPSO mendapatkan hasil sebesar 51,226%
2. Normalisasi atribut dataset berkontribusi untuk dapat meningkatkan hasil akurasi akhir. Dengan menggabungkan atribut dataset tersebut dari 11 atribut menjadi 6 atribut terbukti dapat meningkatkan tingkat akurasi akhir
3. Untuk penelitian selanjutnya agar dapat meningkatkan hasil akhir dari algoritma C5.0 ini dapat digunakan bersamaan dengan metode boost seperti adaboost bersamaan dengan algoritma optimasi yang lain.

DAFTAR PUSTAKA

- [1] T. L. MacKay, N. Bard, M. Bowling, and D. C. Hodgins, "Do pokers players know how good they are? Accuracy of poker skill estimation in online and offline players," *Comput. Human Behav.*, vol. 31, no. 1, pp. 419–424, 2014.
- [2] B. Zaman, K. Geurden, R. De Cock, B. De Schutter, and V. Vanden Abeele, "Motivation profiles of online Poker players and the role of interface preferences: A laddering study among amateur and (semi-) professionals," *Comput. Human Behav.*, vol. 39, pp. 154–164, 2014.
- [3] H. Qing, "Formal verification for winning strategy of chess game," in *2017 International Workshop on Complex Systems and Networks (IWCSN)*, 2017, pp. 55–58.
- [4] S. Li, H. Ma, and M. Ding, "Design and implementation of winning decision algorithm in Hex game system," *Proc. 30th Chinese Control Decis. Conf. CCDC 2018*, pp. 5909–5912, 2018.
- [5] G. Ambekar, T. Chikane, S. Sheth, A. Sable, and K. Ghag, "Anticipation of winning probability in poker using data mining," in *2015 International Conference on Computer, Communication and Control (IC4)*, 2015, vol. 151, pp. 1–6.
- [6] R. Radziukas, R. Maskeliūnas, and R. Damaševičius, "Prediction of poker moves using sequential model and tensorflow," *Commun. Comput. Inf. Sci.*, vol. 1078 CCIS, pp. 516–525, 2019.
- [7] A. Heiberg, "Using Bayesian networks to model a poker player," in *AAAI Workshop - Technical Report*, 2013.
- [8] K. Fan, Y. Ren, and Z. Yan, "Application of Birth Defect Prediction Model Based on C5.0 Decision Tree Algorithm," *2018 IEEE Int. Conf. Internet Things IEEE Green Comput. Commun. IEEE Cyber, Phys. Soc. Comput. IEEE Smart Data*, pp. 1349–1354, 2018.
- [9] S. Mehta and D. Shukla, "Optimization of C5.0 classifier using Bayesian theory," *IEEE Int. Conf. Comput. Commun. Control. IC4 2015*, 2016.
- [10] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 - International Conference on Neural Networks*, 1995, vol. 4, no. 2, pp. 1942–1948.
- [11] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360)*, 1998, vol. 49, no. 5, pp. 69–73.

- [12] C. Qi, A. Fourie, and Q. Chen, "Neural network and particle swarm optimization for predicting the unconfined compressive strength of cemented paste backfill," *Constr. Build. Mater.*, vol. 159, pp. 473–478, 2018.
- [13] M. Anshori, F. Mar'i, M. W. Alauddin, and F. A. Bachtiar, "Prediction Result of Dota 2 Games Using Improved SVM Classifier Based on Particle Swarm Optimization," *3rd Int. Conf. Sustain. Inf. Eng. Technol. SIET 2018 - Proc.*, pp. 121–126, 2018.
- [14] I. Soltani, M. Sarvi, and F. Salahian, "Various Types of Particle Swarm Optimization-based Methods for Harmonic Reduction of Cascade Multilevel Inverters for renewable energy sources," *Int. J. Innov. Appl. Stud.*, vol. 2, no. 4, pp. 671–681, 2013.
- [15] D. T. Larose and C. D. Larose, *Discovering Knowledge in Data*. 2014.
- [16] X. Wu and V. Kumar, *The Top Ten Algorithms in Data Mining*, vol. 53, no. 9. 2009.
- [17] K. P. Wirdhaningsih, D. E. Ratnawati, U. B. Malang, D. Mining, and D. Tree, "Penerapan algoritma decision tree c5.0 untuk peramalan forex," *Doro J.*, vol. 2, no. 8, pp. 1–6, 2013.
- [18] Y. He, W. J. Ma, and J. P. Zhang, "The Parameters Selection of PSO Algorithm influencing On performance of Fault Diagnosis," *MATEC Web Conf.*, vol. 63, no. 2016, p. 02019, 2016.



AUTHOR GUIDELINES

1. Manuscript criteria

- The article has never been published or in the submission process on other publications.
- Submitted articles could be original research articles or technical notes.
- The similarity score from plagiarism checker software such as Turnitin is 20% maximum.
- For December 2021 publication onwards, Ultimatics : Jurnal Teknik Informatika will be receiving and publishing manuscripts written in English only.

2. Manuscript format

- Article been type in Microsoft Word version 2007 or later.
- Article been typed with 1 line spacing on an A4 paper size (21 cm x 29,7 cm), top-left margin are 3 cm and bottom-right margin are 2 cm, and Times New Roman's font type.
- Article should be prepared according to the following author guidelines in this [template](#). Article contain of minimum 3500 words.
- References contain of minimum 15 references (primary references) from reputable journals/conferences

3. Organization of submitted article

The organization of the submitted article consists of Title, Abstract, Index Terms, Introduction, Method, Result and Discussion, Conclusion, Appendix (if any), Acknowledgment (if any), and References.

- Title
The maximum words count on the title is 12 words (including the subtitle if available)
- Abstract
Abstract consists of 150-250 words. The abstract should contain logical argumentation of the research taken, problem-solving methodology, research results, and a brief conclusion.
- Index terms
A list in alphabetical order in between 4 to 6 words or short phrases separated by a semicolon (;), excluding words used in the title and chosen carefully to reflect the precise content of the paper.
- Introduction
Introduction commonly contains the background, purpose of the research,

problem identification, research methodology, and state of the art conducted by the authors which describe implicitly.

- Method
Include sufficient details for the work to be repeated. Where specific equipment and materials are named, the manufacturer's details (name, city and country) should be given so that readers can trace specifications by contacting the manufacturer. Where commercially available software has been used, details of the supplier should be given in brackets or the reference given in full in the reference list.
- Results and Discussion
State the results of experimental or modeling work, drawing attention to important details in tables and figures, and discuss them intensively by comparing and/or citing other references.
- Conclusion
Explicitly describes the research's results been taken. Future works or suggestion could be explained after it
- Appendix and acknowledgment, if available, could be placed after Conclusion.
- All citations in the article should be written on References consecutively based on its' appearance order in the article using Mendeley (recommendation). The typing format will be in the same format as the IEEE journals and transaction format.

4. Reviewing of Manuscripts

Every submitted paper is independently and blindly reviewed by at least two peer-reviewers. The decision for publication, amendment, or rejection is based upon their reports/recommendations. If two or more reviewers consider a manuscript unsuitable for publication in this journal, a statement explaining the basis for the decision will be sent to the authors within six months of the submission date.

5. Revision of Manuscripts

Manuscripts sent back to the authors for revision should be returned to the editor without delay (maximum of two weeks). Revised manuscripts can be sent to the editorial office through the same online system. Revised manuscripts returned later than one month will be considered as new submissions.

6. Editing References

- **Periodicals**
J.K. Author, "Name of paper," Abbrev. Title of Periodical, vol. x, no. x, pp. xxx-xxx, Sept. 2013.
- **Book**
J.K. Author, "Title of chapter in the book," in Title of His Published Book, xth ed. City of Publisher, Country or Nation: Abbrev. Of Publisher, year, ch. x, sec. x, pp xxx-xxx.
- **Report**
J.K. Author, "Title of report," Abbrev. Name of Co., City of Co., Abbrev. State, Rep. xxx, year.
- **Handbook**
Name of Manual/ Handbook, x ed., Abbrev. Name of Co., City of Co., Abbrev. State, year, pp. xxx-xxx.
- **Published Conference Proceedings**
J.K. Author, "Title of paper," in Unabbreviated Name of Conf., City of Conf., Abbrev. State (if given), year, pp. xxx-xxx.
- **Papers Presented at Conferences**
J.K. Author, "Title of paper," presented at the Unabbrev. Name of Conf., City of Conf., Abbrev. State, year.
- **Patents**
J.K. Author, "Title of patent," US. Patent xxxxxxxx, Abbrev. 01 January 2014.
- **Theses and Dissertations**
J.K. Author, "Title of thesis," M.Sc. thesis, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year. J.K. Author, "Title of dissertation," Ph.D. dissertation, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.
- **Unpublished**
J.K. Author, "Title of paper," unpublished.
J.K. Author, "Title of paper," Abbrev. Title of Journal, in press.
- **On-line Sources**
J.K. Author. (year, month day). Title (edition) [Type of medium]. Available: [http://www.\(URL\)](http://www.(URL)) J.K. Author. (year, month). Title. Journal [Type of medium]. volume(issue), pp. if given. Available: [http://www.\(URL\)](http://www.(URL)) Note: type of medium could be online media, CD-ROM, USB, etc.

7. Editorial Adress

Universitas Multimedia Nusantara
Jl. Scientia Boulevard, Gading Serpong
Tangerang, Banten, 15811
Email: ultimatics@umn.ac.id

Paper Title

Subtitle (if needed)

Author 1 Name¹, Author 2 Name², Author 3 Name²

¹Line 1 (of affiliation): dept. name of organization, organization name, City, Country
Line 2: e-mail address if desired

²Line 1 (of affiliation): dept. name of organization, organization name, City, Country
Line 2: e-mail address if desired

Accepted on mmmmm dd, yyyy

Approved on mmmmm dd, yyyy

Abstract—This electronic document is a “live” template which you can use on preparing your ULTIMATICS paper. Use this document as a template if you are using Microsoft Word 2007 or later. Otherwise, use this document as an instruction set. Do not use symbol, special characters, or Math in Paper Title and Abstract. Do not cite references in the abstract.

Index Terms—enter key words or phrases in alphabetical order, separated by semicolon (;)

I. INTRODUCTION

This template, modified in MS Word 2007 and saved as a Word 97-2003 document, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. Margins, column widths, line spacing, and type styles are built-in here. The authors must make sure that their paper has fulfilled all the formatting stated here.

Introduction commonly contains the background, purpose of the research, problem identification, and research methodology conducted by the authors which been describe implicitly. Except for Introduction and Conclusion, other chapter’s title must be explicitly represent the content of the chapter.

II. EASE OF USE

A. *Selecting a Template*

First, confirm that you have the correct template for your paper size. This template is for ULTIMATICS. It has been tailored for output on the A4 paper size.

B. *Maintaining the Integrity of the Specifications*

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them.

III. PREPARE YOUR PAPER BEFORE STYLING

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not add any kind of

pagination anywhere in the paper. Please take note of the following items when proofreading spelling and grammar.

A. *Abbreviations and Acronyms*

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Abbreviations that incorporate periods should not have spaces: write “C.N.R.S.,” not “C. N. R. S.” Do not use abbreviations in the title or heads unless they are unavoidable.

B. *Units*

- Use either SI (MKS) or CGS as primary units (SI units are encouraged).
- Do not mix complete spellings and abbreviations of units: “Wb/m²” or “webers per square meter,” not “webers/m².” Spell units when they appear in text: “...a few henries,” not “...a few H.”
- Use a zero before decimal points: “0.25,” not “.25.”

C. *Equations*

The equations are an exception to the prescribed specifications of this template. You will need to determine whether or not your equation should be typed using either the Times New Roman or the Symbol font (please no other font). To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Number the equations consecutively. Equation numbers, within parentheses, are to position flush right, as in (1), using a right tab stop.

$$\int_0^{r_2} F(r, \phi) dr d\phi = [\sigma r_2 / (2\mu_0)] \quad (1)$$

Note that the equation is centered using a center tab stop. Be sure that the symbols in your equation have been defined before or immediately following the

equation. Use “(1),” not “Eq. (1)” or “equation (1),” except at the beginning of a sentence: “Equation (1) is ...”

D. Some Common Mistakes

- The word “data” is plural, not singular.
- The subscript for the permeability of vacuum μ_0 , and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o.”
- In American English, commas, semi-/colons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an “inset,” not an “insert.” The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
- Do not use the word “essentially” to mean “approximately” or “effectively.”
- In your paper title, if the words “that uses” can accurately replace the word using, capitalize the “u”; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones “affect” and “effect,” “complement” and “compliment,” “discreet” and “discrete,” “principal” and “principle.”
- Do not confuse “imply” and “infer.”
- The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the “et” in the Latin abbreviation “et al.”
- The abbreviation “i.e.” means “that is,” and the abbreviation “e.g.” means “for example.”

IV. USING THE TEMPLATE

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention as below

ULTIMATICS_firstAuthorName_paperTitle.

In this newly created file, highlight all of the contents and import your prepared text file. You are

now ready to style your paper. Please take note on the following items.

A. Authors and Affiliations

The template is designed so that author affiliations are not repeated each time for multiple authors of the same affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

B. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include ACKNOWLEDGMENTS and REFERENCES, and for these, the correct style to use is “Heading 5.”

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles, named “Heading 1,” “Heading 2,” “Heading 3,” and “Heading 4”, are prescribed.

C. Figures and Tables

Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1,” even at the beginning of a sentence.

TABLE I. TABLE STYLES

Table Head	Table Column Head		
	Table column subhead	Subhead	Subhead
copy	More table copy		

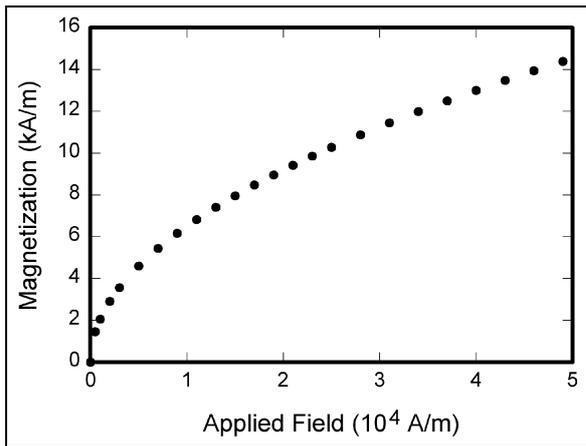


Fig. 1. Example of a figure caption

V. CONCLUSION

A conclusion section is not required. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

APPENDIX

Appendixes, if needed, appear before the acknowledgment.

ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in American English is without an “e” after the “g.” Use the singular heading even if you have many acknowledgments. Avoid expressions such as “One of us (S.B.A.) would like to thank” Instead, write “F. A. Author thanks” You could also state the sponsor and financial support acknowledgments here.

REFERENCES

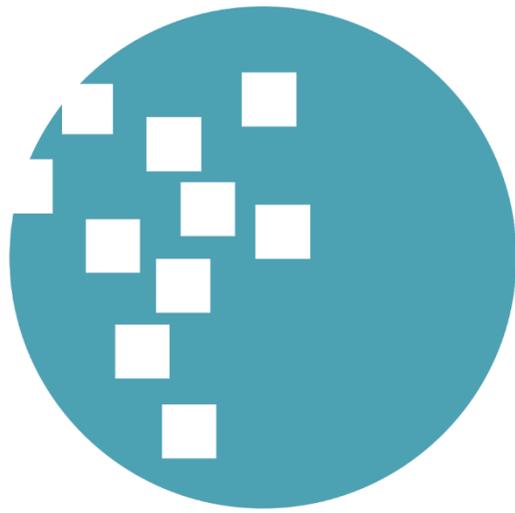
The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

- [1] G. Eason, B. Noble, and I.N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529-551, April 1955. (*references*)
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.
- [3] I.S. Jacobs and C.P. Bean, “Fine particles, thin films and exchange anisotropy,” in *Magnetism*, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.
- [4] K. Elissa, “Title of paper if known,” unpublished.
- [5] R. Nicole, “Title of paper with only first word capitalized,” *J. Name Stand. Abbrev.*, in press.
- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetism Japan, p. 301, 1982].
- [7] M. Young, *The Technical Writer’s Handbook*. Mill Valley, CA: University Science, 1989.



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

ISSN 2085-4552



9 772085 455006



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

PRESS

Universitas Multimedia Nusantara
Scientia Garden Jl. Boulevard Gading Serpong, Tangerang
Telp. (021) 5422 0808 | Fax. (021) 5422 0800