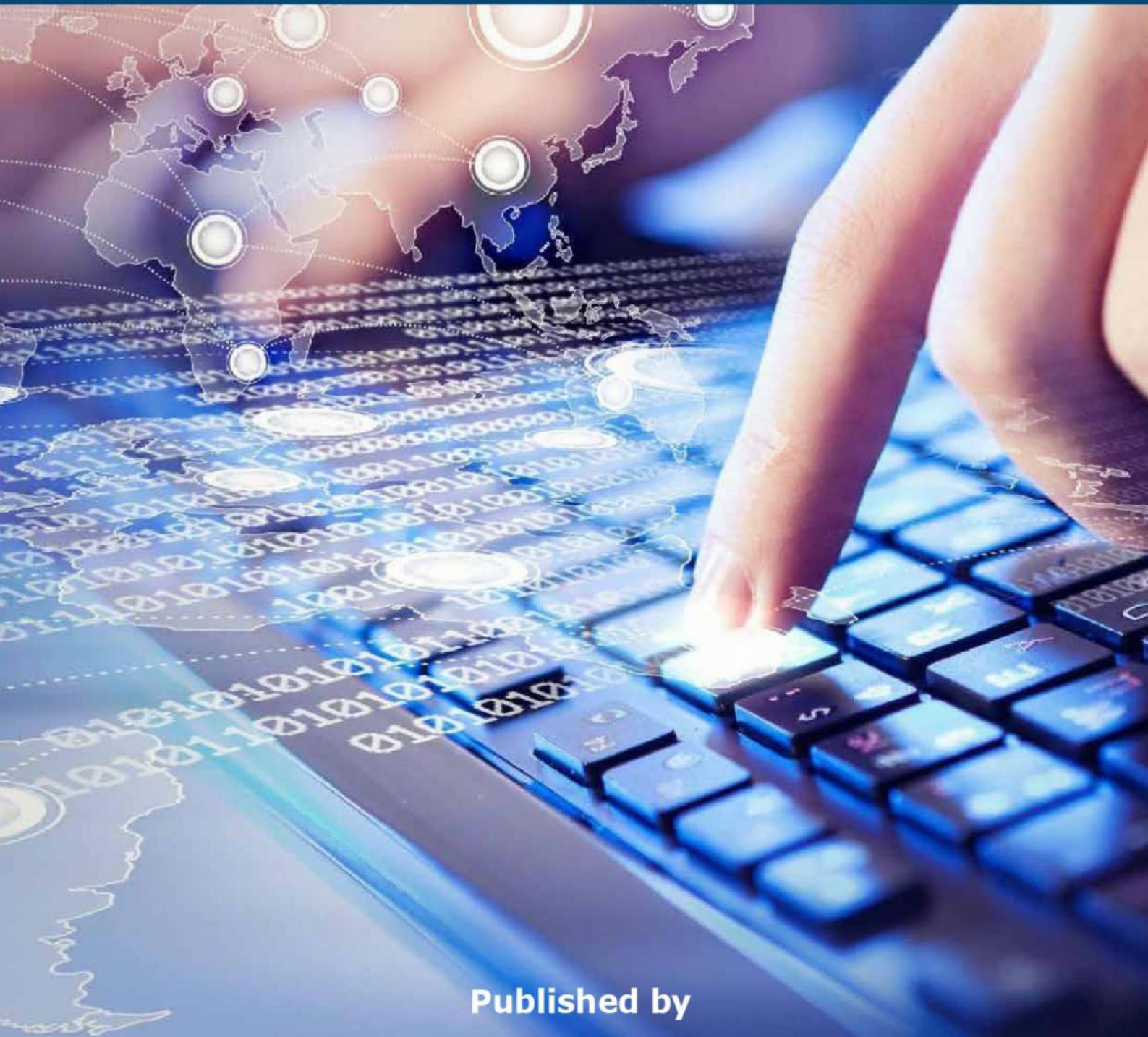


ISSN 2477-0043
E-ISSN 2460-7908

Jurnal Teknologi Terpadu

Volume 9 No. 1, Juli 2023



Published by
LPPM STT TERPADU NURUL FIKRI

Jurnal Teknologi Terpadu

Jurnal Teknologi Terpadu memuat jurnal ilmiah di bidang Ilmu Komputer, Sistem Informasi dan Teknik Informatika. Jurnal Teknologi Terpadu diterbitkan oleh LPPM STT Nurul Fikri dengan periode dua kali dalam setahun, yakni pada bulan Juli dan Desember. Jurnal Teknologi Terpadu telah terakreditasi nasional Sinta 4 sesuai dengan Surat Keputusan No. 5/E/KPT/2022 tanggal 7 Desember 2022 yang dikeluarkan oleh Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi.

Ketua Penyunting (Editor-in-chief)

Drs. Rusmanto, M.M.,
Sistem Informasi,
STT Terpadu Nurul Fikri

Anggota Penyunting (Managing Editor)

Yekti Wirani, S.T., M.T.I.,
Sistem Informasi,
STT Terpadu Nurul Fikri

Dewan Penyunting (Editorial Board Member)

Dr. Yan Riyanto,
Lembaga Ilmu Pengetahuan Indonesia
Indonesia

Dr. Lukman Rosyidi, S.T., M.M., M.T.,
Teknik Informatika,
STT Terpadu Nurul Fikri

Amalia Rahmah, S.T., M.T.,
Teknik Informatika,
STT Terpadu Nurul Fikri

Dr. Sigit Puspito Wigati Jarot,
Teknik Informatika,
STT Terpadu Nurul Fikri

Pristi Sukmasetya, S.Komp., M.kom.,
Universitas Muhammadiyah Magelang
Indonesia

Mitra Bestari (Reviewer)

Indra Hermawan, S.Kom., M.Kom.,
Politeknik Negeri Jakarta, Indonesia

Sirojul Munir, S.Si., M.Kom.,
STT Terpadu Nurul Fikri, Indonesia

Tirsa Ninia Lina, S.Kom., M.Cs.,
Universitas Victory Sorong, Indonesia

Oman Somantri, S.Kom., M.Kom.,
Politeknik Negeri Cilacap, Indonesia

Kelik Sussolaikah, S.Kom., M.Kom.,
Universitas PGRI Madiun, Indonesia

Condro Kartiko, S.Kom., M.T.I.,
Institut Teknologi Telkom Purwokerto,
Indonesia

Matheus Supriyanto Rumetna, S.Kom., M.Cs.,
Universitas Victory Sorong, Indonesia

Tifani Nabarian, S.Kom., M.T.I.,
STT Terpadu Nurul Fikri, Indonesia

Ninik Sri Lestari, S.T., M.Kom.,
STT Mandala, Indonesia

Afif Zuhri Arfianto, S.T, M.T.,
Politeknik Perkapalan Negeri Surabaya,
Indonesia

Edy Victor Haryanto, M.Kom.,
Universitas Potensi Utama, Indonesia

Rismayani, S.Kom., M.T.,
STMIK Dipanegara Makassar, Indonesia

Arnisa Stefanie, S.T., M.T.,
Universitas Singaperbangsa Karawang,
Indonesia

Taufik Hidayat, S.Kom., M.T.,
Universitas Wiralodra, Indonesia

Rahmad Hidayat, S.T., M.T.,
Sekolah Tinggi Teknologi Mandala,
Indonesia

Ahmad Rio Adriansyah, S.Si., M.Si.
STT Terpadu Nurul Fikri, Indonesia

Joko Kuswanto, M.Kom.
Universitas Baturaja, Indonesia

Penyunting Pelaksana (Assistant Editors)

Nurul Janah, S.IIP, M.Hum.,
Sistem Informasi,
STT Terpadu Nurul Fikri

Muh Syaiful Romadhon, S.Kom.,
Sistem Informasi,
STT Terpadu Nurul Fikri

Jurnal Teknologi Terpadu telah terindeks oleh Google Scholar, Index Copernicus International, Garuda, Neliti, dan Sinta. Tanggung jawab isi artikel berada di penulis bukan pada penerbit atau editor.

Diterbitkan oleh:

LPPM STT Terpadu Nurul Fikri

Alamat Redaksi dan Distribusi:

Kampus B STT Terpadu Nurul Fikri Lantai 3

Jl. Lenteng Agung Raya 20, Jakarta Selatan, DKI Jakarta, 12640

Telp. 021 – 786 3191 dan WhatsApp. 0851 7444 3360

Email: journal@nurulfikri.ac.id

Website: <https://journal.nurulfikri.ac.id/index.php/jtt/> dan lppm.nurulfikri.ac.id

Daftar Isi

Kombinasi Linier Target Data untuk Regresi Multitarget menggunakan <i>Principal Component Analysis</i>	01
Yonathan Purbo Santosa	
Sistem Klasifikasi Karakter Kepribadian Siswa Sekolah Dasar berdasarkan Tipologi <i>Hippocrates-Galenus</i> menggunakan Metode Naïve Bayes	10
Muhammad Sabri, Dedy Kasriadi, Irsal, Suci Ramadhani Arifin	
Sistem Informasi Pelayanan Surat Menyurat di Kelurahan Desa Sriwijaya Lampung Tengah	15
Andronias Siregar, Arief Satriansyah, Rachmat Hidayat, Maya Septa Wijaya	
<i>Progressive Web Apps</i>: Pengembangan dan Studi Penerimaan pada Mahasiswa Indonesia menggunakan Scrum dan UTAUT	22
Herman, Frederick	
Perbandingan Model <i>Machine Learning</i> pada Klasifikasi Tumor Otak menggunakan Fitur <i>Discrete Cosine Transform</i>	29
Simeon Yuda Prasetyo, Ghinaa Zain Nabiilah	
Analisis <i>Text Mining</i> Klasifikasi Kegiatan Keluarga menggunakan Orange dengan Metode Naïve Bayes	35
Arsya Fathiarahma, Apriade Voutama, Taufik Ridwan, Nono Heryana	
Analisis Sentimen Ulasan Aplikasi MyPertamina pada <i>Google Play Store</i> menggunakan Algoritma NBC	42
Rihan Maulana, Apriade Voutama, Taufik Ridwan	
Pengembangan Aplikasi <i>Emoticon Recognition</i> dan <i>Facial Recognition</i> menggunakan Algoritma <i>Local Binary Pattern Histogram (LBPH)</i> dan <i>Convolutional Neural Network (CNN)</i>	49
Haeruddin, Herman, Patrick Pratama Hendri	
<i>Design System</i> pada Perancangan Antar Muka Perangkat Lunak Sistem Akses Digital	56
Apriansyah Rizqi Setiawan, Marsani Asfi, Agus Seviana, Sudadi Pranata, Willy Eka Septian	
Pengaruh Pemasaran Media Sosial terhadap Keterlibatan Pelanggan (Survei pada Pengguna Halodoc di Indonesia)	65
Alif Ridha Ramadhani, Mochamad Ardan Fauzi, Muhammad Mufti Abdullah, Syti Sarah Maesaroh, Oding Herdiana	



KOMBINASI LINIER TARGET DATA UNTUK REGRESI MULTITARGET MENGGUNAKAN *PRINCIPAL COMPONENT ANALYSIS*

Yonathan Purbo Santosa¹

¹ Teknik Informatika, Universitas Katolik Soegijapranata
Semarang, Jawa Tengah, Indonesia 50275

yonathansantosa@unika.ac.id

Abstract

Linear regression is a method to predict numbers, a dependent variable (output) based on some independent variables (inputs). The problem with regression is that some data does not fall into linear problems. Based on this problem, RLC was invented to randomly find a correlation between output by projecting the data into the higher dimension. Unfortunately, RLC does not provide ways to inverse the projection, resulting in poor performance results. On top of that, projecting the data into a higher dimension will increase the learning algorithm complexity. Consequently, PCA can solve the problems by projecting the target data into a lower dimension while leaving possibilities for inverse transformation. This research was implemented with the help of the sci-kit-learn library to create and train the regression model and transform the dataset using Python programming language. As a result, for 12 datasets, augmentation using PCA achieved lower error in 7 datasets than RLC, averaging at 0.3270 for augmentation using PCA and 0.4003 for augmentation using RLC.

Keywords: Dimension reduction, linear regression, multidimension regression, multitarget regression, PCA

Abstrak

Regresi linier adalah metode untuk memprediksi sebuah nilai (variabel dependen) berdasarkan beberapa *input* (variabel independen). Permasalahan pada regresi linier adalah beberapa data tidak termasuk ke dalam kategori linier. Sebuah metode bernama RLC diciptakan untuk menemukan korelasi antara data output dengan cara memproyeksikan data ke dalam dimensi yang lebih tinggi. Sayangnya, metode RLC tidak dapat diinvers transformasinya. Selain itu, dengan memproyeksikan data ke dimensi yang lebih tinggi akan menambah kompleksitas dari algoritma pembelajaran. Oleh karena itu, PCA akan digunakan untuk memecahkan masalah ini dengan cara memproyeksikan data ke dimensi yang lebih rendah sembari mempertahankan kemampuan untuk melakukan invers proyeksi. Penelitian ini diimplementasikan dengan bantuan *library scikit-learn* untuk membuat model regresi dan transformasi data dengan menggunakan bahasa pemrograman Python. Hasilnya, untuk 12 dataset, metode augmentasi PCA mampu mendapatkan nilai *error* yang lebih rendah dalam 7 dataset dibandingkan dengan RLC dengan rata-rata nilai *error* 0.3270 untuk metode augmentasi PCA dan 0.4003 untuk metode augmentasi RLC.

Kata kunci: PCA, reduksi dimensi, regresi linier, regresi multidimensi, regresi multitarget

1. PENDAHULUAN

Pada analisis statistik, ada metode yang dapat digunakan untuk memprediksi suatu variabel dependen (*output*) berdasarkan beberapa data yang diberikan sebagai variabel independen (*input*) yang dikenal sebagai metode regresi. Regresi mencari dan menghitung korelasi antara *output* berdasarkan semua variabel *input* untuk mendefinisikan suatu fungsi yang akan memetakan *input* ke *output* [1]. Kita hanya dapat melakukan fungsi estimasi yang akan memetakan *input* ke *output* dengan menemukan deviasi minimum antara fungsi dan variabel dependen karena data

yang dikumpulkan mungkin memiliki *noise*. Dalam kasus linier, korelasi antara variabel dependen dan parameter fungsi dapat diselesaikan dengan fungsi linier seperti fungsi garis, yang biasanya dapat diselesaikan dengan menggunakan estimasi *least square* [2].

Pada masalah yang lebih kompleks, yang terdiri dari *input* dan *output* multidimensi, sering kali tidak termasuk dalam kategori yang dapat diselesaikan secara linier sehingga tidak dapat diselesaikan dengan menggunakan estimasi *least square*. Terlebih, metode regresi hanya mampu

menghasilkan satu nilai prediksi sehingga untuk kasus variabel dependen multidimensi diperlukan beberapa model regresi. Padahal banyak sekali kasus dalam dunia industri yang memerlukan metode prediksi untuk variabel dependen dengan multidimensi [3], [4], terutama dimasa industri 4.0 yang menyebabkan data dengan dimensi yang banyak adalah hal yang sangat wajar [5]–[7]. Selain itu pada bidang kesehatan, model regresi juga terus menjadi kebutuhan utama dalam kasus prediksi suatu penyakit [8]. Didorong oleh kebutuhan untuk memecahkan masalah regresi non-linier dan multidimensi di dunia industri, peneliti di seluruh dunia untuk berusaha untuk menggunakan metode penyelesaian data multidimensi yang non-linier dengan model *machine learning* yang lebih kuat, seperti jaringan syaraf tiruan [9], *deep regression* [10], [11], *random linear target combination* [12], *variational autoencoder regression* [13], *support vector regression* [14], dan lainnya.

Tingginya kebutuhan untuk melakukan komputasi yang lebih kompleks, kemampuan komputer dalam memproses data juga harus ikut meningkat, terutama dalam permasalahan *machine learning* [15]. Tak lain halnya kebutuhan untuk memecahkan permasalahan regresi multidimensi secara akurat pun ikut bertambah seiring dengan kebutuhan industri terutama dalam kasus *big data regression* [16]. Padahal semakin besar dimensi sebuah data, akan semakin besar pula probabilitas untuk terjadi kesalahan dalam pengukuran dan pelabelan pada data [17]. Berdasarkan permasalahan di atas, metode untuk membuat data lebih mudah dipelajari banyak penelitian yang cenderung melakukan pemrosesan pada data ketimbang pada algoritma pembelajaran.

Untuk menyelesaikan permasalahan regresi multidimensi, metode regresi yang sering digunakan adalah dengan menggunakan regresi *single target* (ST) untuk kemudian digabung menjadi sebuah model yang lebih kompleks menggunakan metode *ensemble* seperti pada penelitian Boye dkk. [18] untuk melakukan estimasi harga unit rumah dan berhasil memperoleh nilai standar *error residual* 0.027 satuan. Herawati dkk. [19] membandingkan beberapa metode regresi kembangan untuk menyelesaikan masalah multi-kolinier dalam data dengan cara mengurangi kolinieritas dari variabel independen. Hasilnya menunjukkan *principal component regression* adalah metode dengan nilai *average mean squared error* terendah yaitu 0.0014 dibanding *ridge regression* yang memperoleh nilai *error* 0.4099 [19]. Hal ini menunjukkan dengan mengurangi kolinieritas dari variabel independen, dapat meningkatkan akurasi dari model regresi.

Penelitian yang dilakukan oleh Tsoumakas dkk. [12] memiliki metode pendekatan yang lain, yaitu bahwa kolinieritas dari variabel independen akan berkurang dengan cara melakukan transformasi terhadap variabel dependen. Dalam penelitian tersebut, variabel dependen ditransformasi secara *random* dengan hanya memilih beberapa atribut pada variabel dependen dan melakukan

transformasi ke dimensi yang lebih tinggi. Metode tersebut diberi nama *random linear target combination* (RLC) pada regresi ST. Meskipun menggunakan kombinasi *random*, hasil dari penelitian tersebut menunjukkan performansi yang lebih baik (dinilai dari nilai *error*) dibandingkan dengan *state-of-the-art* sebelumnya yaitu *multi-objective random forest algorithm* (MORF) yang hanya lebih unggul dalam 4 *dataset* dari 12 *dataset* yang diuji coba .

Pada metode augmentasi RLC, data target ditransformasi ke dimensi yang lebih tinggi hingga 30 kali lipat lebih besar, sehingga dapat menyebabkan berkurangnya stabilitas dari sebuah model *regressor* yang sering dikenal dengan istilah *curse of dimensionality* di dalam *data mining* maupun *machine learning* [20]. Selain itu semakin tinggi dimensi sebuah data, semakin tinggi pula jumlah *noise* yang terdapat pada data ketika data tersebut dibuat atau dikumpulkan [21]. Hal ini dianggap sebuah permasalahan penting karena dalam proses regresi, keakuratan dari prediksi sangat penting.

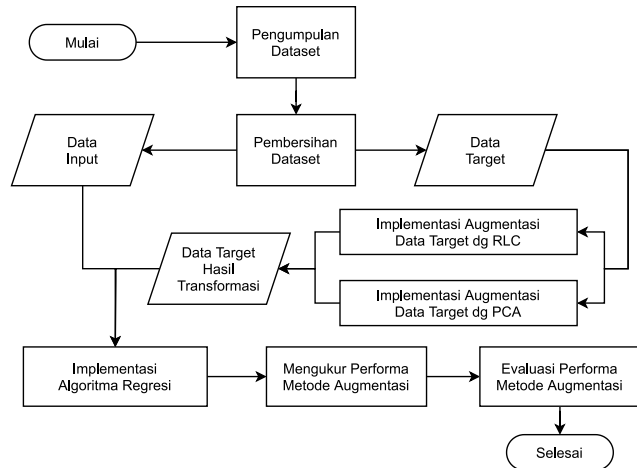
Tidak hanya permasalahan dimensi, proses transformasi data target dengan metode RLC menggunakan perkalian matriks yang tidak dapat dilakukan transformasi inversnya karena matriks transformasi tersebut tidak memenuhi syarat matriks yang dapat dihitung inversnya. Hal ini menyebabkan hasil dari prediksi tidak dapat digunakan untuk menghasilkan nilai prediksi dengan satuan yang sama dengan data. Sebagai contoh, model regresi untuk harga rumah harus bisa menghasilkan prediksi berupa harga rumah [18]. Oleh sebab itu, untuk menghitung matriks invers matriks tersebut harus menggunakan metode estimasi *Moore-Penrose Psuedoinverse* (MPPI). Dikarenakan metode transformasi tersebut adalah sebuah estimasi, keakuratan dari model *regresi* akan tergantung dari keakuratan proses estimasi invers matriks tersebut. Hal ini didukung oleh penelitian Górecki dan Łuczak [22], yang dalam pengujian menggunakan 15 *dataset*, algoritma genetika sebagai metode generalisasi dari *psuedoinverse* lebih unggul dalam 14 *dataset* dibandingkan dengan hanya menggunakan MPPI. Sayangnya, algoritma genetika berjalan sangat lambat untuk mencapai titik optimal jika *dataset* terlalu besar dikarenakan kompleksitas komputasi dari algoritma genetika [22], [23]. Hal ini mendorong untuk mencari metode transformasi data target yang dapat dihitung nilai inversnya tanpa metode estimasi guna memperoleh nilai *error* terendah tanpa mengorbankan komputasi yang kompleks.

Keterbatasan pada penelitian terdahulu dalam hal mengatasi permasalahan *curse of dimensionality* dan penggunaan matriks *psuedoinverse* yang dapat menyebabkan kurang akuratnya hasil prediksi, mendorong penelitian ini untuk dilakukan dengan tujuan memberikan usulan metode pengganti yang bekerja dengan cara mengurangi dimensi melalui kombinasi lainnya guna mencegah *curse of dimensionality* tanpa menggunakan metode estimasi *psuedoinverse* dalam melakukan proses transformasi.

Dengan demikian, hasil dari regresi diduga akan memiliki nilai *error* yang lebih rendah dalam rangka mendukung industri 4.0 dalam melakukan regresi linier terutama untuk data dengan variabel dependen yang memiliki multidimensi.

2. METODE PENELITIAN

Pada bagian menjelaskan metode untuk melakukan analisa terhadap efek dari perubahan dimensi pada data target, baik dari penelitian sebelumnya yaitu RLC dan metode usulan yang digunakan pada penelitian ini. Alur penelitian ini dirancang seperti pada Gambar 1 yang akan dijabarkan secara lebih mendalam pada bagian berikutnya.



Gambar 1. Diagram Alir Penelitian

Uji coba dan evaluasi perbandingan metode transformasi data serta implementasi model regresi linier dibuat menggunakan bahasa pemrograman Python versi 3.7 dengan bantuan *library machine learning* Scikit-learn versi 1.1.3 [24], dan Numpy versi 1.23 [25]. *Library* Scikit-learn memuat *Gradient Tree Boosting* untuk melakukan ensambel regresi linier serta transformasi data yang dibantu oleh *library* Numpy. Program diimplementasikan menggunakan komputer dengan spesifikasi Windows 10, Intel Core I7-6200U, RAM 16GB, dan *harddisk* 500GB.

2.1 Pengumpulan Dataset

Pengujian metode transformasi membutuhkan *dataset* yang digunakan sebagai bahan percobaan dalam memperoleh hasil pengukuran metode augmentasi yang digunakan. Jika mengacu kepada penelitian yang dilakukan Tsoumakas dkk. [12], setidaknya 12 *dataset* yang sama juga digunakan pada penelitian ini seperti yang dijabarkan pada Tabel 1. Hal ini bertujuan untuk mendapatkan komparasi yang sama jika dibandingkan dengan penelitian terdahulu [12]. Deskripsi lengkap serta metode pengumpulan data dari masing-masing *dataset* dijabarkan secara lengkap pada penelitian yang menyediakan *dataset* tersebut [12], [26]–[29]. *Dataset* dapat didapatkan pada sumber penyedia *dataset* dan dapat diunduh secara bebas. Pada Tabel 1, dijabarkan bahwa masing-masing *dataset* memiliki data target dengan ukuran dimensi q . *Dataset* tersebut kemudian diunduh dan diproses

lebih lanjut menggunakan bahasa pemrograman Python dan *library* Numpy agar dapat dilakukan transformasi data dan untuk melakukan proses regresi linier melalui program. Proses transformasi data target dari masing-masing *dataset* dilakukan menggunakan metode PCA dan RLC yang akan digunakan untuk menguji keefektifan kedua algoritma tersebut dalam menangani permasalahan regresi multitarget.

Tabel 1. *Dataset* dengan Deskripsi

Nama	Singkatan	$ D $	q
<i>Airline Ticket Price</i> 1 [28]	atp1d	337	6
<i>Airline Ticket Price</i> 2 [28]	atp7d	296	6
<i>Occupational Employment Survey</i> 1 [28]	oes10	334	16
<i>Occupational Employment Survey</i> 2 [28]	oes97	403	16
<i>River Flow</i> 1 [28]	rf1	4165/5065	8
<i>River Flow</i> 2 [28]	rf2	4165/5065	8
<i>Supply Chain Management</i> 1 [28]	sf1978	8145/1658	16
<i>Supply Chain Management</i> 2 [28]	sf1969	7463/1503	16
<i>Electrical Discharge Machining</i> [29]	edm	154	2
<i>Solar Flare</i> 1 [26]	scm1d	323	3
<i>Solar Flare</i> 2 [26]	scm20d	1066	3
<i>Water Quality</i> [27]	wq	1060	14

Penjelasan singkat mengenai *dataset* pada Tabel 1 akan dijelaskan pada bagian di bawah. *Dataset Airline Ticket Price* adalah *dataset* yang berisi harga tiket pesawat untuk 1 hari ke depan (atp1d) dan harga tiket terendah untuk 7 hari ke depan (atp7d) untuk maskapai Delta, maskapai Continental, maskapai Airtran, dan maskapai United [28]. *Dataset Occupational Employment Survey* terdiri dari data penerimaan sebagai karyawan yang dilakukan oleh Badan Statistik Kepegawaian Amerika Serikat (*US Bureau of Labor Statistics*) pada tahun 2010 (oes10) dan pada tahun 1997 (oes97) [28]. *Dataset River Flow* dibuat dalam rangka melakukan prediksi terhadap jaringan sungai pada 8 kota yang berbeda untuk 48 jam ke depan [28]. *Dataset River Flow* ada 2 jenis data, *dataset* pertama berisi observasi aliran (rf1) sedangkan *dataset* kedua ditambahkan juga data mengenai kelembapan udara [28]. *Dataset Supply Chain Management* didapatkan dari *Trading Agent Competition in Supply Chain Management* (TAC SCM) pada tahun 2010 [28]. *Dataset* tersebut terdiri dari data dengan prediksi untuk 1 hari ke depan (scm1d) dan data dengan prediksi untuk 20 hari ke depan (scm20d) [28]. *Dataset Electrical Discharge Machining* berisi data yang mereproduksi perilaku manusia yang bertugas sebagai petugas dalam mengoperasikan sebuah mesin [29]. *Dataset Solar Flare* berisi tentang tipe *solar flare* yang diamati selama kurun waktu 24 jam [26]. Terdapat dua jenis data, pertama data pengamatan pada tahun 1969 dan kedua tahun 1978 [26]. *Dataset Water Quality* berisi tentang kondisi fisik dan kimia pada kualitas air di sungai Slovenia [27]. Semua *dataset* tersebut diunduh dalam format *csv* untuk dapat dibaca oleh *library* Numpy menggunakan bahasa pemrograman Python.

2.2 Pembersihan *Dataset*

Data yang sudah diunduh harus diproses terlebih dahulu untuk membersihkan data yang tidak lengkap dan data yang tidak dapat diolah oleh regresi seperti data yang berupa teks dan kategorial. Data yang termasuk ke dalam kategori tersebut di atas, diubah ke dalam bentuk angka yang merepresentasikan tipe data yang ada pada masing-masing *dataset*. Selain itu dikarenakan data mengandung satuan nilai yang berbeda-beda, data juga akan dilakukan pemrosesan awal *MinMaxScaler* seperti yang dijabarkan pada persamaan (1) dan (2). Hal ini bertujuan untuk mempermudah proses perbandingan dari kedua metode augmentasi PCA dan RLC agar dapat dilakukan dengan mudah.

$$X_{\sigma} = \frac{x - \min(X)}{\max(X) - \min(X)} \quad (1)$$

$$\hat{x} = X_{\sigma} * (\max(X) - \min(X)) + \min(X) \quad (2)$$

2.3 Implementasi Augmentasi Data Target menggunakan RLC

Dalam melakukan augmentasi data target dua metode akan digunakan dalam penelitian ini. Metode yang pertama adalah metode RLC yang menjadi landasan dalam penelitian ini dan metode PCA sebagai alternatif yang akan menggantikan RLC. Metode RLC sendiri berusaha menemukan korelasi linier antar data target dengan cara melakukan inisialisasi sebuah matriks kombinasi C secara *random* dengan tahapan sebagai berikut:

1. Pilih ukuran dimensi data target yang baru r . Pada penelitian sebelumnya, parameter r yang digunakan adalah 500 [12].
2. Pilih jumlah data target dari data asli yang akan dikorelasikan k . Pada penelitian sebelumnya, parameter k berkisar antara 2 hingga q .
3. Buat matriks kosong C dengan ukuran $q \times r$
4. Untuk masing-masing kolom pada C , pilih k buah baris dan isikan nilai dengan nilai *random* yang diambil dari distribusi uniform dengan rentang 0 hingga 1 inklusif.

Berdasarkan langkah-langkah di atas dapat diilustrasikan pada Gambar 2. Sebagai contoh terdapat 3 buah baris data dengan 2 data target Y yang diilustrasikan pada Gambar 2(a). Kemudian dibuatlah sebuah matriks C yang diisi dengan nilai *random*, diilustrasikan dengan Gambar 2(b). Hasil perkalian antara data target Y dengan matriks C kemudian akan digunakan untuk melakukan proses pelatihan model regresi linier yang diilustrasikan pada Gambar 2(c). Proses kombinasi target yang dijabarkan pada dilakukan hanya untuk *training* data, sehingga akan dihasilkan sebuah matriks baru \hat{Y} yang akan digunakan sebagai data target dalam melakukan proses regresi menggunakan data target \hat{Y} seperti pada Gambar 2(c).

Tsoumakas dkk. [12] tidak melakukan proses invers untuk mengubah \hat{Y} menjadi Y kembali.

1	2
4	1
3	4

(a)

0.2	0	0	0.5	0.3	0.3
0.1	0.4	0.1	0	0.2	0

(b)

0.4	0.8	0.2	0.5	0.9	0.3
0.9	0.4	0.1	2	1.4	1.2
1	1.6	0.4	1.5	1.7	0.9

(c)

Gambar 2. (a) Data target asli Y , (b) Matrix Kombinasi C , (c) Data target dari hasil kombinasi.

Pada penelitian ini akan dilakukan proses invers menggunakan MPPI terhadap matriks C untuk menemukan C^+ . Asumsi tersebut didasarkan bahwa ada sebuah matriks C^+ yang memenuhi persamaan berikut:

$$C^+ = (C^T C)^{-1} C^T \quad (3)$$

$$C^+ \approx C^{-1} \quad (4)$$

Dari persamaan (3), dapat dihitung kembali transformasi inversnya agar dapat dilihat nilai prediksi yang mengacu kepada *dataset*. Proses invers transformasi tersebut dijabarkan pada persamaan (5).

$$Y = \hat{Y} \times C^+ \quad (5)$$

Semua proses yang telah dijabarkan diimplementasikan menggunakan bahasa pemrograman Python dan *library Numpy* dan *Scikit-learn*. Setelah diimplementasikan, program dapat digunakan untuk melakukan augmentasi pada data yang sudah dibersihkan.

2.4 Implementasi Augmentasi Data Target menggunakan PCA

Metode kedua yang merupakan metode usulan dari penelitian ini menggunakan metode transformasi PCA. Proses transformasi PCA membutuhkan perhitungan matriks C yang didapatkan dari *singular value decomposition* (SVD). Proses SVD menghasilkan *principal components* (PCs) yang digunakan untuk mencari data yang dengan jumlah variasi tertinggi. Dengan melakukan limitasi terhadap sebagian data yang memiliki variasi tertinggi, data dapat diproyeksikan ke dimensi yang lebih rendah dengan langkah-langkah sebagai berikut:

1. Pilih ukuran dimensi data target yang baru k , yang berada pada rentang antara 2 hingga q
2. Hitung menggunakan SVD untuk mendapatkan *eigenvector* C dan *eigenvalue* Σ
3. Pilih sub set dari Σ dengan ukuran k secara diagonal untuk membuat matriks ukuran urut mulai dari PCs yang paling signifikan
4. Lakukan transformasi berdasarkan matriks C

Persamaan (6) adalah persamaan *singular value decomposition* (SVD) untuk matriks kovarian S dari data target Y dan dapat dihitung matriks dekomposisinya yang merupakan matriks transformasi C menggunakan SVD. Matriks C pada persamaan (6) hingga (8) adalah matriks ortogonal yang berisi eigenvector dari matriks kovarian S , Σ adalah matriks diagonal yang berisikan *singular value* dari S , \hat{Y} adalah hasil transformasi dari matriks Y menggunakan matriks C .

$$Y^T Y = S = U \Sigma C^T \quad (6)$$

$$\hat{Y} = Y \times C \quad (7)$$

$$Y = \hat{Y} \times C^T \quad (8)$$

Dikarenakan matriks C adalah matriks ortogonal, invers dari matriks C sama dengan *transpose* dari matriks C . Dengan demikian, proses transformasi dapat memenuhi sistem persamaan linier (7) dan (8). Hal ini menunjukkan proses transformasi dapat dicari transformasi inversnya tanpa metode estimasi seperti MPPI.

2.5 Implementasi Model Regresi

Pengujian metode augmentasi data dilakukan dengan model regresi linear yang melakukan regresi linier terhadap masing-masing variabel target (ST singkatan dari satu target) dengan menggunakan algoritma *decision tree regressor* yang dilatih menggunakan metode ensemble *gradient boosting*. Untuk mengakomodasi seluruh variabel target, maka akan diperlukan sebanyak k buah ST untuk masing-masing metode augmentasi untuk setiap dataset. Parameter dari model regresi yang digunakan mengacu pada parameter yang didefinisikan dalam penelitian sebelumnya [12]. Model regresi diimplementasikan menggunakan bahasa pemrograman Python dengan *library Scikit-learn*. Selain parameter yang dijabarkan di dalam Tabel 2, semua parameter menggunakan nilai *default* dari implementasi *library Scikit-learn* [24]. Selain itu, untuk parameter *random generator* akan menggunakan angka 0.

Tabel 2. Parameter *Gradient Boosting Tree Regressor*

Parameter	Nilai
Jumlah maksimum <i>leaf node</i>	4
Jumlah <i>decision tree regressor</i>	100

2.6 Pengukuran Performa Metode Augmentasi

Dalam rangka mengetahui performa dari masing-masing metode, *cost function Average Relative Root Mean Squared Error* (aRRMSE) yang telah didefinisikan di dalam penelitian yang menjadi landasan penelitian ini oleh Tsoumakas dkk. [12] akan digunakan sebagai alat ukur. Sebelum perhitungan aRRMSE dilakukan, proses invers dari augmentasi RLC dan PCA, serta *MinMaxScaler* akan dilakukan. aRRMSE sendiri dijabarkan dalam persamaan (9) dan (10).

$$RRMSE = \sqrt{\frac{\sum_{(x,y) \in D_{test}} (h(x)_j - y_j)^2}{\sum_{(x,y) \in D_{test}} (\bar{y}_j - y_j)^2}} \quad (9)$$

$$aRRMSE(h, D_{test}) = \frac{1}{k} \sum_{j=1}^k RRMSE \quad (10)$$

Pada persamaan (9), $h(x)_j$ adalah *output* dari regresi linier untuk seluruh data *input* x untuk data target y_j sedangkan \bar{y}_j nilai rata-rata untuk data target ke- j , k adalah jumlah dimensi dari data target. Dengan kata lain, fungsi aRRMSE, masing-masing *output* dari regresi linier akan dihitung nilai deviasinya terhadap data target dengan mengacu pada nilai rata-rata masing-masing data target. Hal ini bertujuan untuk melakukan standarisasi performa melalui rata-rata dari data target yang tidak homogen setelah proses *training* model ensemble [12].

2.7 Evaluasi Performa Kedua Metode Augmentasi

Di akhir dari proses pelatihan model selesai, hasil dari perhitungan nilai *error* menggunakan *cost function* dari kedua belas *dataset* terhadap data tes akan dibandingkan dengan cara mencari nilai terendah untuk masing-masing *dataset*. Metode dengan nilai *error* yang lebih rendah adalah metode yang dianggap lebih berhasil dalam melakukan regresi linier. Jumlahan keberhasilan dari masing-masing metode untuk semua *dataset* kemudian dibandingkan untuk memperoleh kesimpulan.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Uji Coba Metode *Random Linear Target Combination*

Hasil uji coba metode RLC dijabarkan dalam Tabel 3. Nilai k pada Tabel 3 merupakan jumlahan data target yang dikorelasikan secara linier dan *random*. Pada metode RLC dimensi dari data target akan menjadi konstan yaitu 500 dimensi. Hasil uji coba ini akan menjadi tolok ukur dari uji coba metode PCA. Untuk mempermudah interpretasi dari hasil uji coba, nilai *error* yang merupakan nilai minimum untuk masing-masing *dataset* dicetak tebal dan digarisbawah.

Mengacu kepada hasil uji coba yang dilakukan oleh Tsoumakas dkk. [12] pada penelitian sebelumnya dengan *library* yang berbeda pada Tabel 4, jika dibandingkan dengan Tabel 3 dapat terlihat bahwa hasil uji coba yang dilakukan memiliki nilai *error* yang lebih rendah. Perbedaan ini diduga dikarenakan perbedaan *library* dan *tools* yang digunakan yang diduga menyebabkan proses *random* yang berbeda. Oleh karena hasil dari implementasi di dalam penelitian ini memiliki rata-rata nilai *error* yang lebih rendah, maka hasil uji coba dengan metode RLC pada penelitian ini akan digunakan sebagai acuan.

Tabel 3. Hasil Uji Coba Metode RLC

k	atp1d	atp7d	edm	oes10	oes97	rf1	rf2	sf1978	sf1969	wq	scm1d	scm20d
2	0.4137	0.4078	0.6703	0.5649	0.5247	0.1710	0.1830	1.1424	0.5188	0.1021	0.1140	0.0998
3	0.4016	0.3794		0.5726	0.5255	0.1687	0.1802	1.1268	0.5377	0.0982	0.1139	0.0986
4	0.4032	0.3671		0.5789	0.5351	0.1656	0.1784			0.0938	0.1168	0.0973
5	0.3965	0.3659		0.5907	0.5284	0.1665	0.1790			0.0934	0.1171	0.0977
6	0.3900	0.3660		0.5822	0.5345	0.1638	0.1739			0.0956	0.1194	0.0977
7				0.5876	0.5408	0.1666	0.1741			0.0966	0.1199	0.0989
8				0.5961	0.5385	0.1654	0.1738			0.0950	0.1215	0.0986
9				0.5871	0.5415					0.1009	0.1226	0.0988
10				0.5876	0.5413					0.1006	0.1221	0.0986
11				0.5980	0.5413					0.1032		0.1007
12				0.6029	0.5414					0.1049		0.1000
13				0.5912	0.5404					0.1109		0.1003
14				0.5957	0.5401					0.1172		0.1011
15				0.5985	0.5438							0.1012
16				0.6038	0.5424							0.1011
MIN	0.3900	0.3659	0.6703	0.5649	0.5247	0.1638	0.1738	1.1268	0.5188	0.0934	0.1139	0.0973

Tabel 4. Hasil Uji Coba RLC dari Penelitian Sebelumnya oleh Tsoumakas dkk. [12]

k	atp1d	atp7d	edm	oes10	oes97	rf1	rf2	sf1978	sf1969	wq	scm1d	scm20d
2	0.3842	0.4614	0.6996	0.5026	0.5593	0.7265	0.7036	1.2312	1.5746	0.9100	0.4572	0.7469
3	0.3840	0.4653		0.5084	0.5588	0.7878	0.7584	1.2172	1.5675	0.9080	0.4610	0.7467
4	0.3884	0.4796		0.5232	0.5730	0.8204	0.7922			0.9085	0.4663	0.7472
5	0.3952	0.4917		0.5359	0.5837	0.8584	0.8327			0.9086	0.4699	0.7477
6	0.4022	0.5029		0.5472	0.5889	0.8515	0.8257			0.9089	0.4775	0.7490
7				0.5551	0.5958	0.8446	0.8106			0.9090	0.4820	0.7513
8				0.5734	0.6076	0.8868	0.8655			0.9107	0.4855	0.7536
9				0.5911	0.6153					0.9122	0.4889	0.7548
10				0.6031	0.6229					0.9128	0.4932	0.7537
11				0.6154	0.6348					0.9150	0.4978	0.7573
12				0.6285	0.6449					0.9163	0.5020	0.7571
13				0.6354	0.6590					0.9188	0.5057	0.7619
14				0.6428	0.6682					0.9217	0.5133	0.7640
15				0.6525	0.6860						0.5155	0.7681
16				0.6652	0.6916						0.5218	0.7704
MIN	0.3840	0.4614	0.6996	0.5026	0.5588	0.7268	0.7036	1.2172	1.5675	0.9080	0.4572	0.7467

3.2 Hasil Uji Coba Metode *Principal Component Analysis*

Hasil uji coba metode PCA disajikan pada Tabel 5, masing-masing *dataset* memiliki jumlah uji coba yang berbeda karena perbedaan jumlah data target seperti yang dijabarkan pada Tabel 1. Tergantung dari jumlah data targetnya (q), untuk masing-masing proyeksi ke dimensi yang lebih rendah (k) dari data target akan digunakan untuk melatih *decision tree regressor* dengan *gradient boosting*. Sebagai contoh, untuk *dataset* atp1d yang memiliki $q = 6$, akan dilakukan uji coba dengan memproyeksikan data target

dengan dimensi $k \in \{2, 3, 4, 5, 6\}$. Dari hasil pengujian tersebut akan diambil parameter k mana yang memiliki nilai *error* paling rendah untuk tiap-tiap *dataset*. Dapat diamati juga pada Tabel 5 pola dari masing-masing *dataset* yang memiliki nilai *error* yang paling rendah, yaitu pada nilai k yang cenderung lebih kecil (berkisar antara 2 hingga 5), sehingga semakin besar dimensi yang dimiliki oleh data target dari masing-masing *dataset*, semakin besar pula jumlah reduksi dimensi yang terjadi untuk mendapatkan nilai *error* yang lebih rendah pada pelatihan model regresi ketimbang metode RLC yang mentransformasi data targetnya ke dalam 500 dimensi.

Tabel 5. Hasil Uji Coba PCA

k	atp1d	atp7d	edm	oes10	oes97	rf1	rf2	sf1978	sf1969	wq	scm1d	scm20d
2	0.3921	0.3755	0.8731	0.3849	0.3934	0.1801	0.1895	0.5368	0.3372	0.0804	0.1196	0.0870
3	0.3855	0.4681		0.4521	0.3842	0.1807	0.1898	0.5362	0.3907	0.0836	0.1197	0.0872
4	0.3870	0.5049		0.4495	0.4018	0.1741	0.1905			0.0846	0.1198	0.0876
5	0.3890	0.5110		0.4509	0.4137	0.1722	0.1878			0.0855	0.1224	0.0916
6	0.3870	0.5115		0.4486	0.4460	0.1756	0.1900			0.0888	0.1267	0.0959
7				0.4489	0.4522	0.1759	0.1903			0.0929	0.1275	0.0983
8				0.4514	0.4514	0.1759	0.1903			0.0918	0.1356	0.1191
9				0.4442	0.4504					0.0947	0.1347	0.1165
10				0.4444	0.4499					0.0949	0.1360	0.1199
11				0.4449	0.4515					0.0996	0.1393	0.1223
12				0.4451	0.4500					0.0987	0.1397	0.1217
13				0.4449	0.4504					0.0988	0.1397	0.1232
14				0.4457	0.4512					0.0999	0.1434	0.1258
15				0.4455	0.4512						0.1457	0.1275
16				0.4461	0.4512						0.1483	0.1279
MIN	0.3855	0.3755	0.8731	0.3849	0.3842	0.1722	0.1878	0.5362	0.3372	0.0804	0.1196	0.0870

Pada Tabel 6, hasil uji coba untuk masing-masing *dataset* diurutkan mulai dari *dataset* yang memiliki dimensi data target yang paling kecil hingga *dataset* yang memiliki dimensi data target yang paling besar. Dapat terlihat dari hasil uji coba tersebut, persentase reduksi dimensi yang dibutuhkan untuk memperoleh nilai *error* yang kecil. Untuk *dataset* dengan jumlah dimensi yang lebih dari 10, nilai *error* minimum dapat dicapai menggunakan reduksi hingga 88% dari dimensi aslinya.

Tabel 6. Persentase Reduksi Dimensi terhadap Nilai *Error* Minimum

<i>Dataset</i>	k	q	Persentase Reduksi Dimensi
edm	2	2	0%
sf1969	2	3	33%
sf1978	3	3	0%
atp1d	3	6	50%
atp7d	2	6	67%
rf1	5	8	38%
rf2	5	8	38%
wq	2	14	86%
scm1d	2	16	88%
oes97	3	16	81%
scm20d	2	16	88%
oes10	2	16	88%

3.3 Perbandingan RLC dan PCA dalam Regresi Multitarget

Hasil dari uji coba RLC dan PCA dibandingkan dengan cara membandingkan tiap-tiap *dataset* dengan hasil *error* yang didapatkan, hanya hasil *error* terkecil yang akan dibandingkan. Rangkuman hasil *error* tersebut dapat

dijabarkan di dalam Tabel 7 dengan nilai rata-rata keseluruhan uji coba terhadap dua belas *dataset* disajikan pada baris terakhir.

Tabel 7. Perbandingan aRRMSE metode RLC dan PCA

<i>Dataset</i>	RLC	PCA
atp1d	0.39	0.3855
atp7d	0.3659	0.3755
edm	0.6703	0.8731
oes10	0.5649	0.3849
oes97	0.5247	0.3842
rf1	0.1638	0.1722
rf2	0.1738	0.1878
sf1978	1.1268	0.5362
sf1969	0.5188	0.3372
wq	0.0934	0.0804
scm1d	0.1139	0.1196
scm20d	0.0973	0.087
AVG	0.4003	0.327

Dari hasil uji coba tersebut dapat dilihat bahwa metode PCA memiliki nilai *error* yang lebih rendah untuk *dataset* atp1d, oes10, oes97, sf1978, sf1969, wq, dan scm20d. Sehingga jika dihitung jumlah *dataset* berdasarkan metode yang memiliki *error* yang lebih rendah, didapatkan perbandingan RLC : PCA = 5 : 7. Selain perbandingan untuk masing-masing *dataset*, metode RLC dan PCA memiliki hasil rata-rata *error* masing-masing untuk semua *dataset* yaitu 0.4003 dan 0.3270. Dari hasil perbandingan *error* kedua metode tersebut, dapat ditentukan bahwa PCA mampu menghasilkan nilai *error* yang lebih rendah untuk

sebagian besar dari *dataset* yang diuji coba pada penelitian ini sebagai kesimpulan

3.4 Implementasi Metode PCA pada Penelitian Lanjutan

Pada hasil penelitian ini, dapat dibuktikan bahwa transformasi menggunakan metode PCA dapat menghasilkan nilai *error* yang lebih rendah dibandingkan dengan metode RLC. Selain itu, PCA mampu mereduksi jumlah dimensi dari data target. Hal ini tentu dapat membantu mengurangi nilai *error* dalam membuat sebuah model regresi. Dari temuan tersebut, metode PCA dapat digunakan sebagai sebuah metode pemrosesan awal dalam membuat sebuah model regresi, terutama untuk data dengan dimensi data target yang cukup besar karena dapat mengurangi nilai *error* secara lebih optimal.

4. KESIMPULAN

Dalam kasus regresi multitarget kedua metode dapat menghasilkan rata-rata *error* yang lebih rendah dari 0.8, hal ini menunjukkan simpangan dari hasil regresi tidak akan lebih dari 1 satuan untuk masing-masing *dataset* untuk kedua metode tersebut. Perubahan dari metode transformasi data ke dimensi yang lebih rendah hingga 88% total dimensi aslinya menggunakan metode transformasi PCA daripada transformasi data ke dimensi yang lebih tinggi dapat menghindarkan model regresi dari *curse of dimensionality*. Selain itu, proses invers dari metode transformasi menggunakan PCA tidak menggunakan metode estimasi *psuedoinverse*. Kedua hal tersebut menunjukkan lebih unggulnya metode transformasi PCA dibanding dengan metode transformasi RLC yang dapat ditunjukkan dari lebih rendahnya rata-rata nilai *error* ketika dibandingkan serta untuk masing-masing *dataset*, metode PCA lebih unggul dalam 7 *dataset* dibandingkan dengan RLC yang hanya lebih unggul dalam 5 *dataset*.

Berdasarkan simpulan yang sudah dijabarkan di atas, metode PCA dapat digunakan dalam meningkatkan performa dari model *regresi* khususnya untuk *dataset* yang memiliki target multidimensi. Saran penelitian lanjutan dari hasil penelitian ini adalah dengan menguji coba kedua metode pada model regresi yang lebih kuat seperti *neural network*, *deep regression*, *support vector regression*, dan model regresi lainnya yang mampu melakukan regresi untuk data non-linier dengan lebih baik.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih sebesar-besarnya atas kesempatan yang diberikan oleh LPPM Universitas Katolik Soegijapranata yang telah mendukung secara finansial untuk dapat melakukan penelitian dan menghasilkan tulisan ini.

DAFTAR PUSTAKA

[1] C. Wallisch *et al.*, "Review of guidance papers on regression modeling in statistical series of medical

journals," *PLoS One*, vol. 17, no. 1, p. e0262918, Jan. 2022, doi: 10.1371/JOURNAL.PONE.0262918.

- [2] Anila. M and G. Pradeepini, "Least Square Regression for Prediction Problems in Machine Learning using R," *International Journal of Engineering & Technology*, vol. 7, no. 3.12, pp. 960–962, Jul. 2018, doi: 10.14419/IJET.V7I13.12.17612.
- [3] T. Nabarian, M. Aris Ganiardi, and R. F. Malik, "Implementasi Metode Hibrid Fuzzy C-Means dan Fuzzy Swarm untuk Pengelompokan Data Benang Perusahaan Tekstil," *Jurnal Teknologi Terpadu*, vol. 6, no. 1, pp. 39–45, Jul. 2020, doi: 10.54914/JTT.V6I1.247.
- [4] Carudin, "Pemanfaatan Data Transaksi untuk Dasar membangun Strategi berdasarkan Karakteristik Pelanggan dengan Algoritma K-Means Clustering dan Model RFM," *Jurnal Teknologi Terpadu*, vol. 7, no. 1, pp. 7–14, Jul. 2021, doi: 10.54914/JTT.V7I1.318.
- [5] J. N. Hussain, "High dimensional data challenges in estimating multiple linear regression," *J Phys Conf Ser*, vol. 1591, no. 1, p. 12035, 2020, doi: 10.1088/1742-6596/1591/1/012035.
- [6] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Comput Sci*, vol. 2, no. 3, pp. 1–21, May 2021, doi: 10.1007/S42979-021-00592-X/FIGURES/11.
- [7] S. Jameel and S. Schockaert, "Modeling context words as regions: An ordinal regression approach to word embedding," *CoNLL 2017 - 21st Conference on Computational Natural Language Learning, Proceedings*, pp. 123–133, 2017, doi: 10.18653/V1/K17-1014.
- [8] S. A. T. al Azhima, D. Darmawan, N. F. A. Hakim, I. Kustiawan, M. al Qibtiya, and N. S. Syafei, "Hybrid Machine Learning Model untuk memprediksi Penyakit Jantung dengan Metode Logistic Regression dan Random Forest," *Jurnal Teknologi Terpadu*, vol. 8, no. 1, pp. 40–46, Jul. 2022, doi: 10.54914/JTT.V8I1.539.
- [9] M. Bataineh and T. Marler, "Neural network for regression problems with reduced training sets," *Neural Networks*, vol. 95, pp. 1–9, 2017, doi: <https://doi.org/10.1016/j.neunet.2017.07.018>.
- [10] S. Lathuilière, P. Mesejo, X. Alameda-Pineda, and R. Horaud, "A Comprehensive Analysis of Deep Regression," *IEEE Trans Pattern Anal Mach*

- Intell*, vol. 42, no. 9, pp. 2065–2081, 2020, doi: 10.1109/TPAMI.2019.2910523.
- [11] D. Rügamer *et al.*, “deepregression: a Flexible Neural Network Framework for Semi-Structured Deep Distributional Regression,” *arXiv:2104.02705 [cs, stat]*, 2021, [Online]. Available: <http://arxiv.org/abs/2104.02705>
- [12] G. Tsoumakas, E. Spyromitros-Xioufis, A. Vrekou, and I. Vlahavas, “Multi-Target Regression via Random Linear Target Combinations,” *arXiv:1404.5065 [cs]*, vol. 8726, pp. 225–240, 2014, doi: 10.1007/978-3-662-44845-8_15.
- [13] Q. Zhao, E. Adeli, N. Honnorat, T. Leng, and K. M. Pohl, “Variational AutoEncoder for Regression: Application to Brain Aging Analysis,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11765 LNCS, pp. 823–831, 2019, doi: 10.1007/978-3-030-32245-8_91/COVER.
- [14] Muthukrishnan R and Maryam Jamila S, “Predictive Modeling Using Support Vector Regression,” *International Journal of Scientific & Technology Research*, vol. 9, no. 2, pp. 4863–4865, Feb. 2020, Accessed: Dec. 06, 2022. [Online]. Available: www.ijstr.org
- [15] K. Berggren *et al.*, “Roadmap on emerging hardware and technology for machine learning,” *Nanotechnology*, vol. 32, no. 1, p. 012002, Oct. 2020, doi: 10.1088/1361-6528/ABA70F.
- [16] W. Chiang, X. Liu, T. Zhang, and B. Yang, “A Study of Exact Ridge Regression for Big Data,” *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, pp. 3821–3830, Jan. 2019, doi: 10.1109/BIGDATA.2018.8622274.
- [17] D. Xu, Y. Shi, I. W. Tsang, Y.-S. Ong, C. Gong, and X. Shen, “A Survey on Multi-output Learning,” Jan. 2019, doi: 10.48550/arxiv.1901.00248.
- [18] P. Boye, D. Mireku-Gyimah, and C. A. Okpoti, “Multiple Linear Regression Model for Estimating the Price of a Housing Unit,” *Ghana Mining Journal*, vol. 17, no. 2, pp. 66–77, 2017, doi: 10.4314/gm.v17i2.9.
- [19] N. Herawati, K. Nisa, E. Setiawan, N. Nusyirwan, and T. Tiryono, “Regularized multiple regression methods to deal with severe multicollinearity,” *Int J Stat Appl*, vol. 8, no. 4, pp. 167–172, 2018.
- [20] O. Eguasa, E. Edionwe, and J. I. Mbegbu, “Local Linear Regression and the problem of dimensionality: a remedial strategy via a new locally adaptive bandwidths selector,” <https://doi.org/10.1080/02664763.2022.2026895>, 2022, doi: 10.1080/02664763.2022.2026895.
- [21] Y. Xu, S. Balakrishnan, A. Singh, and A. Dubrawski, “Regression with Comparisons: Escaping the Curse of Dimensionality with Ordinal Information,” *Journal of Machine Learning Research*, vol. 21, no. 162, pp. 1–54, 2020, [Online]. Available: <http://jmlr.org/papers/v21/19-505.html>
- [22] T. Górecki and M. Łuczak, “Stacked Regression With a Generalization of the Moore-Penrose Pseudoinverse,” *Statistics in Transition New Series*, vol. 18, no. 3, pp. 443–458, 2017, doi: 10.21307/stattrans-2016-080.
- [23] S. Katoch, S. S. Chauhan, and V. Kumar, “A review on genetic algorithm: past, present, and future,” *Multimed Tools Appl*, vol. 80, no. 5, pp. 8091–8126, Feb. 2021, doi: 10.1007/S11042-020-10139-6/FIGURES/8.
- [24] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [25] C. R. Harris *et al.*, “Array programming with NumPy,” *Nature* 2020 585:7825, vol. 585, no. 7825, pp. 357–362, Sep. 2020, doi: 10.1038/s41586-020-2649-2.
- [26] A. Asuncion and D. Newman, “UCI: Machine Learning Repository : Solar Flare Dataset,” 1989. <http://archive.ics.uci.edu/ml/datasets/Solar+Flare> (accessed Jun. 05, 2022).
- [27] S. Džeroski, D. Demsar, and J. Grbović, “Predicting Chemical Parameters of River Water Quality from Bioindicator Data,” *Applied Intelligence*, vol. 13, pp. 7–17, 2000, doi: 10.1023/A:1008323212047.
- [28] E. Spyromitros-Xioufis, G. Tsoumakas, W. Groves, and I. Vlahavas, “Multi-target regression via input space expansion: treating targets as inputs,” *Mach Learn*, vol. 104, no. 1, pp. 55–98, 2016, doi: 10.1007/s10994-016-5546-z.
- [29] A. Karalic and I. Bratko, “First Order Regression,” *Mach Learn*, vol. 26, pp. 147–176, 1997, doi: 10.1023/A:1007365207130.



SISTEM KLASIFIKASI KARAKTER KEPERIBADIAN SISWA SEKOLAH DASAR BERDASARKAN TIPOLOGI *HIPPOCRATES-GALENUS* MENGGUNAKAN METODE NAÏVE BAYES

Muhammad Sabri¹, Dedy Kasriadi², Irsal³, Suci Ramadhani Arifin⁴

^{1,2,3,4} Program Studi Teknik Informatika, Universitas Dipa Makassar
Makassar, Sulawesi Selatan, Indonesia 90245

muhammadsabri1306@gmail.com, dedykasriadi63@gmail.com, irsal@undipa.ac.id, suci.arifin@undipa.ac.id

Abstract

Schools must shape students' personalities through learning activities and programs that can make students become educated human beings with noble characters as the nation's next generation. Teachers as educators need to recognize students' character to develop more effective teaching methods. So far, the guidance process by teachers at SD Inpres 12/79 Palattae has been carried out face-to-face, namely by talking directly to students and carried out without knowing the student's character. As a result, the guidance and counseling that is carried out are sometimes ignored by students. Based on these circumstances, this research was then carried out to build an information system that implements the Naïve Bayes method in elementary school students' personality character classification system based on the Hippocrates-Galenus typology to know the personality types possessed by students. With this system, the teacher can assist in knowing the personality types possessed by students so that it makes it easier to build the mindset, attitude, and behavior of students so that they become positive, good-natured, noble-spirited, and responsible individuals.

Keywords: *character, character classification, hippocrates-galenus, naïve bayes, student character*

Abstrak

Sekolah dituntut untuk membentuk kepribadian siswa melalui kegiatan dan program pembelajaran yang mampu menjadikan siswa menjadi manusia yang terpelajar dan berakhlak mulia sebagai generasi penerus bangsa. Penting bagi guru sebagai pendidik untuk mengenali karakter siswa untuk mengembangkan metode pengajaran yang lebih efektif. Selama ini, proses bimbingan oleh guru di SD Inpres 12/79 Palattae dilaksanakan secara tatap muka yaitu dengan berbicara langsung pada siswa dan dilakukan tanpa mengetahui karakter siswa tersebut. Akibatnya, bimbingan dan konseling yang dilaksanakan kadang-kadang tidak dipedulikan oleh siswa. Berdasarkan keadaan tersebut, penelitian ini kemudian dilaksanakan untuk membangun sistem informasi yang mengimplementasikan metode *Naïve Bayes* dalam sistem klasifikasi karakter kepribadian siswa sekolah dasar berdasarkan tipologi *Hippocrates-Galenus* dengan tujuan untuk mengetahui karakter tipe kepribadian yang dimiliki oleh siswa. Dengan adanya sistem tersebut, Guru BK dapat membantu dalam mengetahui karakter tipe kepribadian yang dimiliki oleh siswa sehingga memudahkan dalam membangun pola pikir, sikap, dan perilaku peserta didik agar menjadi pribadi yang positif, berakhlak karimah, berjiwa luhur, dan bertanggung jawab.

Kata kunci: *hippocrates-galenus, karakter, karakter siswa, klasifikasi karakter, naïve bayes*

1. PENDAHULUAN

Sekolah dituntut untuk mampu membentuk karakter siswa melalui kegiatan pembelajaran serta program yang mampu membentuk karakter siswa di sekolah [1], [2]. Sekolah berperan penting dan memiliki pengaruh yang cukup besar sebagai wadah pembentukan karakter siswa. Para guru juga dituntut untuk mampu menjadi teladan bagi siswa dalam rangka membentuk karakter yang baik [3].

SD Inpres 12/79 Palattae adalah salah satu sekolah yang terletak di Kabupaten Bone Provinsi Sulawesi Selatan yang menjunjung tinggi kepribadian siswa khususnya pembentukan karakter yang berakhlak karimah, berjiwa luhur, dan bertanggung jawab. Dalam kehidupan berbangsa dan bernegara, karakter adalah hal mendasar, kehilangan karakter sama saja dengan kehilangan generasi muda penerus bangsa [4]. Disisi lain, sebagai penggerak dan kekuatan adalah fungsi karakter yang menjaga bangsa ini agar tidak mudah goyah dan terombang-ambing [5].

Selama ini, proses bimbingan oleh guru di SD Inpres 12/79 Palattae dilaksanakan secara tatap muka yaitu dengan berbicara langsung pada siswa dan dilakukan tanpa mengetahui karakter siswa tersebut. Akibatnya, bimbingan dan konseling yang dilaksanakan kadang-kadang tidak dipedulikan oleh siswa. Berdasarkan keadaan tersebut, penelitian ini kemudian dilaksanakan dengan tujuan untuk mengetahui karakter tipe kepribadian yang dimiliki oleh siswa.

Data siswa dikelompokkan berdasarkan data yang diperoleh dari kuesioner siswa melalui pertanyaan yang diberikan sebagai data latih. Data latih akan dimasukkan di *microsoft excel* sebagai data uji, kemudian diimpor ke dalam sistem yang akan tersimpan sebagai *database*, selanjutnya sistem menyediakan beberapa pertanyaan melalui *form website* yang akan disimpan sebagai data uji. Metode *Naïve Bayes* digunakan dalam mengelompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya yang telah diperoleh [6]. Tujuannya adalah melakukan *cluster* pada objek menurut atribut partisi [7]–[9]. Metode ini awalnya ditentukan oleh *cluster* yang akan disusun, kemudian elemen pertama tiap *cluster* dapat dijadikan titik tengah (*centroid*), lalu dilakukan pengulangan prosedur sampai tidak ada lagi objek yang dapat dipindahkan [7], [10], [11].

Beberapa penelitian terkait Tipologi *Hippocrates-Galenus* telah dilakukan. Prawito [12] membuat sebuah desain dengan tujuan mempermudah serta membantu guru BK dalam mencari tahu tentang karakteristik kepribadian setiap siswa sesuai dengan Tipologi *Hippocrates-Galenus*. Hasil yang didapatkan yaitu berupa informasi untuk membantu pihak guru BK mengetahui karakter tipe kepribadian yang dimiliki oleh siswa SMP. Meilana *et al* [13] menerapkan algoritma *Naïve Bayes Classifier* untuk mengklasifikasikan data berdasarkan karakteristik yang sudah di tentukan sebelumnya, yakni *sanguin*, *koleris*, *melankolis*, *plegmatis* untuk menentukan kelas tertinggi yang akan ditunjukkan pada

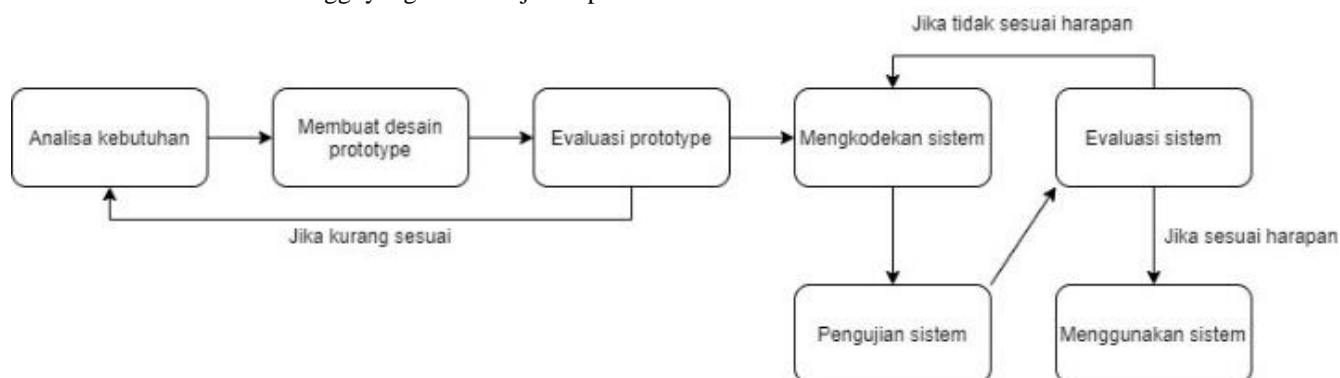
karakteristik kepribadian tersebut dan menyimpulkan bahwa algoritma *Naïve Bayes classifier* mampu mengklasifikasikan dengan baik dan memperoleh hasil akurasi yang cukup tinggi. Nasrul [16] melakukan penelitian melakukan berupa pengembangan sistem informasi pemetaan kompetensi pengajar LP3 STT-NF berbasis web menggunakan *Yii Framework*. Azizah [14] melakukan penelitian korelasional yang bertujuan untuk mengetahui hubungan tipe kepribadian *Hippocrates-Galenus* dengan tingkat kesabaran mahasiswa dalam mengikuti perkuliahan *online*. Hasil penelitian Azizah menunjukkan bahwa tidak terdapat hubungan yang signifikan antara tipe kepribadian *Hippocrates-Galenus* dengan tingkat kesabaran mahasiswa.

Adapun hasil dari penelitian ini, guru BK dapat melihat nilai persentase *Sanguin*, *Kholeris*, *Melankolis*, dan *Plegmatis*. Dengan adanya sistem tersebut, Guru BK dapat membantu dalam mengetahui karakter tipe kepribadian yang dimiliki oleh siswa. Sehingga memudahkan dalam membangun pola pikir, sikap, dan perilaku peserta didik agar menjadi pribadi yang positif, berakhlak karimah, berjiwa luhur, dan bertanggung jawab.

2. METODE PENELITIAN

Penelitian ini adalah penelitian lapangan dengan metode pengambilan data observasi dan wawancara. Penelitian ini dilaksanakan di SD Inpres 12/79 Palattae beralamat di Jalan Jl. Ahmad Yani, Palattae, Kec. Kahu, Kabupaten Bone, Sulawesi Selatan, 92767, Indonesia. Jumlah siswa yang menjadi sampel dalam penelitian ini adalah sebanyak 38 orang. Kuesioner yang digunakan dalam penelitian ini diadopsi dari penelitian sebelumnya yang telah dilakukan oleh Aprilla [15].

Dalam merancang dan mengembangkan sistem, penelitian ini menggunakan metode *Prototype* dengan tahapan sebagai berikut (Gambar 1).



Gambar 1. Metode *Prototype*

Tahap pertama adalah tahap analisa kebutuhan untuk mengidentifikasi kebutuhan dalam perancangan sistem yang akan dibuat. Tahap berikutnya adalah membuat *prototype* atau membuat rancangan sementara dari sistem untuk dipresentasikan kepada pengguna yaitu guru dan

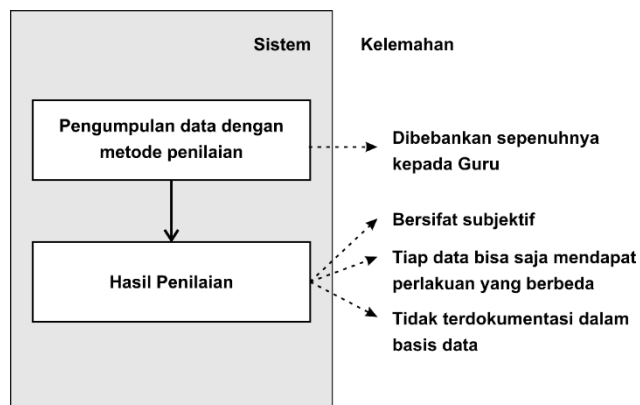
siswa. Setelah membuat *prototype*, kemudian dilaksanakan evaluasi *prototype* dengan tujuan untuk meminta umpan balik dari pengguna terhadap *prototype* yang telah dipresentasikan melalui diskusi secara langsung. Hasil dari diskusi ini akan dipertimbangkan apakah sistem perlu

direvisi atau tidak. Jika tidak maka dapat lanjut ke tahap berikutnya yaitu mengkodekan sistem dengan cara menerjemahkan sistem ke dalam bentuk aplikasi berbasis *website* dengan menggunakan bahasa pemrograman PHP dan DBMS MySQL. Langkah selanjutnya adalah pengujian sistem dengan metode pengujian *black box* diikuti dengan evaluasi sistem untuk mengevaluasi kinerja sistem dan tingkat akurasi dalam menyelesaikan masalah, apakah sesuai dengan harapan pengguna atau tidak. Terakhir adalah implementasi sistem diawali dengan melakukan sosialisasi tentang aplikasi sistem yang telah dibuat.

Dalam proses bimbingan karakter secara khusus oleh penanggung jawab bimbingan konseling (dalam hal ini guru BK) maupun dalam kegiatan belajar-mengajar secara luas, setiap guru sebagai tenaga pengajar membutuhkan informasi mengenai karakter siswanya. Dengan demikian guru diharapkan dapat lebih mengenal siswanya sebagai manusia yang kompleks sehingga dapat melaksanakan dan mengembangkan metode pembelajaran yang relevan dan bersifat personal.

Sampai saat ini, informasi mengenai karakter kepribadian siswa belum disediakan oleh pihak sekolah. Guru yang membutuhkan informasi ini harus melakukan riset mandiri dalam mengumpulkan data-data yang dibutuhkan untuk ditransformasi ke dalam bentuk informasi agar bisa digunakan.

Dari sistem yang digunakan saat ini, ditemukan beberapa kelemahan sebagai berikut (Gambar 2).



Gambar 2. Kelemahan sistem yang sedang berjalan

Kebutuhan sistem sebagai data yang dibutuhkan dalam membangun sistem adalah sebagai berikut (Tabel 1).

Tabel 1. Kebutuhan Sistem

Bahan penelitian	Keterangan
Data Siswa	Terdapat 221 data siswa yang aktif dengan jumlah populasi pada periode tahun 2020-2021 dari kelas 1 sampai 6.
Data Kuesioner	Total 60 pertanyaan dari masing-masing 15 pertanyaan setiap karakteristik

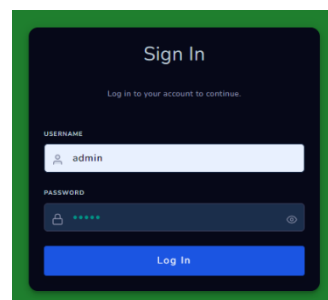
Pada sistem ini, pengguna diklasifikasikan berdasarkan hak akses dari tiap fitur yang disediakan. Jika dideskripsikan dalam bentuk objek di dunia nyata, kategori pengguna sistem dapat dikategorikan sebagai berikut (Tabel 2).

Tabel 2. Kategori Pengguna Sistem

Pengguna	Tugas
Administrator	memiliki akses untuk mengelola data tiap siswa dan manajemen akun pengguna, melihat hasil dari klasifikasi oleh sistem untuk semua siswa, serta memasukkan, memperbaharui, dan menghapus pertanyaan kuesioner.
Guru	memiliki akses dalam mengelola data tiap siswa, mengisi kuesioner dan melihat hasil dari klasifikasi oleh sistem untuk semua siswa.
Siswa	memiliki akses dalam membaca dan menjawab pertanyaan kuesioner, dan melihat hasil dari klasifikasi. Untuk jenis pengguna berupa siswa hanya dapat mengakses hasil dari tesnya sendiri dan tidak dapat mengakses data siswa lain.
Orang tua siswa	mendapatkan laporan hasil klasifikasi melalui pesan <i>Whatsapp</i> oleh guru.

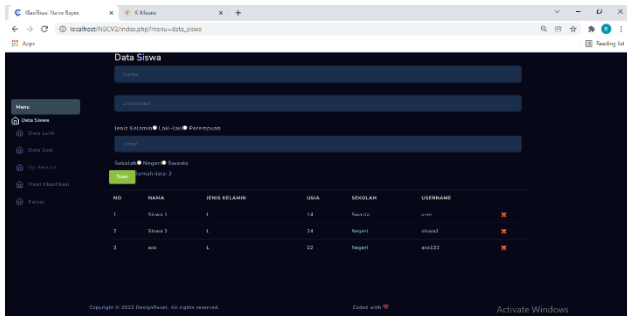
3. HASIL DAN PEMBAHASAN

Implementasi metode *Naïve Bayes* dalam sistem klasifikasi karakter kepribadian siswa sekolah dasar berdasarkan tipologi *Hippocrates-Galenus* disajikan pada Gambar 3 sampai Gambar 8.



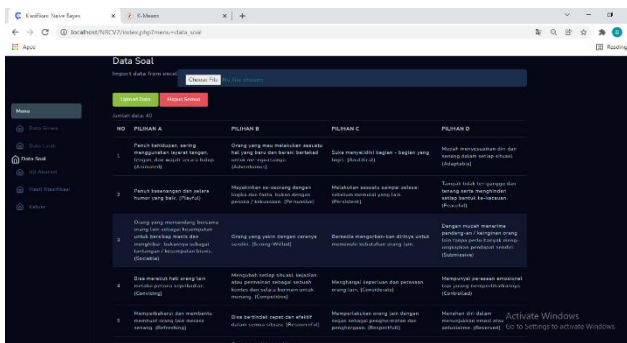
Gambar 3. Tampilan Form Login

Setelah administrator masuk di halaman *website*, admin terlebih dahulu melakukan *login* untuk masuk ke sistem (Gambar 3). Admin akan memasukkan *username* dan *password* pada kolom yang tersedia dan kemudian menekan tombol *login*. Jika *username* dan *password* salah maka pesan kesalahan akan ditampilkan, jika *username* atau *password* benar maka sistem akan meneruskan ke halaman utama (Gambar 4).



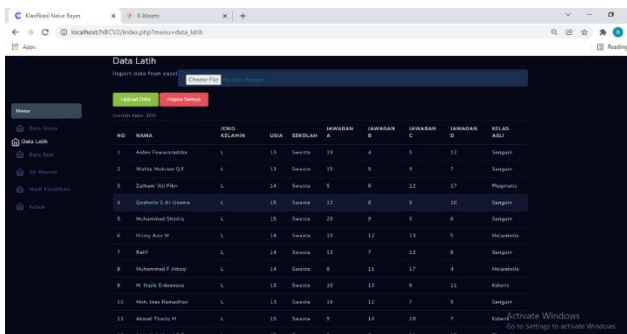
Gambar 4. Tampilan Form Data Siswa

Halaman data siswa (Gambar 4) dapat menampilkan daftar data siswa, dapat menambah data dengan memasukkan data dengan lengkap dan dapat mengubah data yang diinginkan dan dapat menghapus data dari *database*.



Gambar 5. Tampilan Form Data Soal

Tampilan berikutnya adalah *form* data soal (Gambar 5). Pada *form* ini halaman sistem dapat menampilkan daftar data soal, dapat menambah data dengan memasukkan data dengan lengkap dan dapat menghapus data dari *database*.



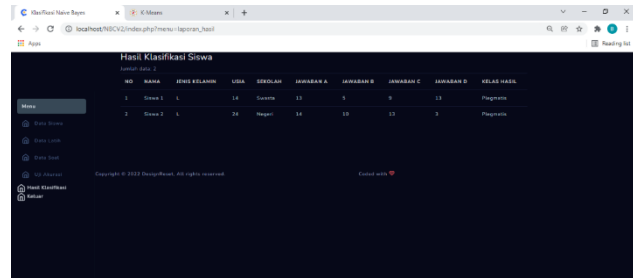
Gambar 6. Tampilan Form Data Latih

Tampilan berikutnya adalah *form* data latih (Gambar 6). Pada *form* ini halaman sistem menampilkan data latih dan dapat menampilkan data jawaban berdasarkan data soal.



Gambar 7. Tampilan Form Klasifikasi Kepribadian

Tampilan berikutnya adalah *form* klasifikasi kepribadian (Gambar 7). Halaman ini menampilkan pertanyaan kuesioner yang dapat diisi oleh siswa untuk disimpan sebagai jawaban kuesioner.



Gambar 8. Tampilan Form Hasil Analisis Naïve Bayes

Tampilan berikutnya adalah *form* klasifikasi kepribadian (Gambar 8). Untuk halaman ini menampilkan daftar data kepribadian siswa dan admin dapat melihat detail nilai kepribadian.

Tabel 3. Rekapitulasi Hasil Pengujian Black Box

No.	Modular	Berhasil
1.	Pengujian Login	✓
2.	Pengujian Form Data Siswa	✓
3.	Pengujian Form Data Soal	✓
4.	Pengujian Form Data latih	✓
5.	Pengujian Form Klasifikasi Siswa	✓
6.	Pengujian Form Data Hasil Analisis Naïve Bayes	✓
Total		17

Berdasarkan dari hasil pengujian yang terlihat pada rekapitulasi hasil pengujian *blackbox* (Tabel 3) menunjukkan tingkat keberhasilan sebanyak 100%. Dengan demikian aplikasi Klasifikasi Kepribadian Siswa ini dapat dikatakan benar dan layak untuk dipergunakan.

4. KESIMPULAN

Algoritma *Naïve Bayes* dapat digunakan untuk melakukan klasifikasi karakteristik kepribadian siswa sekolah dasar menjadi empat kategori dalam tipologi *Hippocrates-Galenus*. Berdasarkan hasil pengujian menunjukkan tingkat keberhasilan sebanyak 100%. Dengan demikian aplikasi Klasifikasi Kepribadian Siswa berbasis *web* ini layak dipergunakan. Dengan sistem yang diterapkan, para guru dimudahkan dalam menemukan informasi karakter

kepribadian siswa dalam aplikasi yang dapat diakses kapan dan di mana saja dikarenakan ketersediaan aplikasi dalam bentuk *web*. Selain digunakan dalam Kegiatan Belajar Mengajar (KMB) setiap harinya, guru juga bisa mengirimkan hasil penilaian karakter kepada orang tua masing-masing siswa.

DAFTAR PUSTAKA

- [1] U. Khasanah and H. Herina, "Membangun Karakter Siswa Melalui Literasi Digital Dalam Menghadapi Pendidikan Abad 21 (Revolusi Industri 4.0)," *Prosiding Seminar Nasional Program Pascasarjana Universitas PGRI Palembang*, vol. 12, no. 01, Art. no. 01, Mar. 2019, Accessed: Jun. 15, 2022. [Online]. Available: <https://jurnal.univpgri-palembang.ac.id/index.php/Prosidingpps/article/view/2662>
- [2] A. Harita, B. Laia, and S. F. L. Zagoto, "Peranan Guru Bimbingan Konseling Dalam Pembentukan Karakter Disiplin Siswa Smp Negeri 3 Onolalu Tahun Pelajaran 2021/2022," *Counseling For All (Jurnal Bimbingan dan Konseling)*, vol. 2, no. 1, Art. no. 1, Mar. 2022.
- [3] E. C. Hendriana and A. Jacobus, "Implementasi Pendidikan Karakter Di Sekolah Melalui Keteladanan Dan Pembiasaan," *JPDI (Jurnal Pendidikan Dasar Indonesia)*, vol. 1, no. 2, pp. 25–29, Oct. 2017, doi: 10.26737/jpdi.v1i2.262.
- [4] D. Setiawan, "Pendidikan Kewarganegaraan Berbasis Karakter melalui Penerapan Pendekatan Pembelajaran Aktif, Kreatif, Efektif dan Menyenangkan," *JUPIIS: Jurnal Pendidikan Ilmu-Ilmu Sosial*, vol. 6, no. 2, Art. no. 2, Jan. 2015, doi: 10.24114/jupiis.v6i2.2285.
- [5] F. Robiansyah, B. Tristyanto, F. Alfarisa, A. Murdaningsih, and A. Hijratunnisa, "Pembinaan Karakter Siswa SD Melalui Kegiatan Ekstrakurikuler Memanah:," *Prosiding Didaktis: Seminar Nasional Pendidikan Dasar*, vol. 4, no. 1, Art. no. 1, Dec. 2019.
- [6] D. Wilandini and P. Purwantoro, "Penerapan Algoritma Naïve Bayes dalam mengklasifikasikan Media Sosial untuk mengamati Trend Kuliner," *Jurnal Teknologi Terpadu*, vol. 8, no. 1, Art. no. 1, Jul. 2022, doi: 10.54914/jtt.v8i1.535.
- [7] E. G. Sihombing, "Klasifikasi Data Mining Pada Rumah Tangga Menurut Provinsi dan Status Kepemilikan Rumah Kontrak/Sewa Menggunakan K-Means Clustering Method," *CESS (Journal of Computer Engineering, System and Science)*, vol. 2, no. 2, Art. no. 2, Jul. 2017, doi: 10.24114/cess.v2i2.6347.
- [8] Z. Nabila, A. R. Isnain, P. Permata, and Z. Abidin, "Analisis Data Mining untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *Jurnal Teknologi dan Sistem Informasi*, vol. 2, no. 2, Art. no. 2, Jul. 2021, doi: 10.33365/jtsi.v2i2.868.
- [9] T. Suprawoto, "Klasifikasi Data Mahasiswa Menggunakan Metode K-Means Untuk Menunjang Pemilihan Strategi Pemasaran," *JIKO (Jurnal Informatika dan Komputer)*, vol. 1, no. 1, Art. no. 1, Feb. 2016, doi: 10.26798/jiko.v1i1.9.
- [10] S. Rustam, "Analisa Clustering Phising Dengan K-Means Dalam Meningkatkan Keamanan Komputer," *ILKOM Jurnal Ilmiah*, vol. 10, no. 2, Art. no. 2, Aug. 2018, doi: 10.33096/ilkom.v10i2.309.175-181.
- [11] U. Ma'rifatin, "Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien Pada Puskesmas Warujayeng," *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, vol. 4, no. 1, Art. no. 1, Dec. 2020, doi: 10.29407/inotek.v4i1.211.
- [12] A. Prawito, "Klasifikasi Karakteristik Kepribadian Manusia Berdasarkan Tipologi Hippocrates–Galenus Menggunakan Metode Naïve-Bayes," *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, vol. 5, no. 1, Art. no. 1, Aug. 2021.
- [13] M. Meilana, Y. Astuti, I. R. Wulandari, I. Sulistyowati, and B. A. Mimartiningtyas, "Algoritma Naive Bayes untuk Mengklasifikasikan Kepribadian Siswa SMP Berdasarkan Tipologi Hippocrates-Galenus," *Sistemasi: Jurnal Sistem Informasi*, vol. 10, no. 2, Art. no. 2, May 2021, doi: 10.32520/stmsi.v10i2.1339.
- [14] N. D. Azizah, "Hubungan Tipe Kepribadian Hippocrates-Galenus dengan Tingkat Kesabaran Mahasiswa dalam Sistem Kuliah Online," *Jurnal Perspektif*, vol. 5, no. 1, Art. no. 1, Jun. 2021, doi: 10.15575/jp.v5i1.116.
- [15] L. Aprilla, "Klasifikasi Karakteristik Kepribadian Manusia Berdasarkan Tipologi Hippocrates-Galenus Menggunakan Metode Naïve Bayes," undergraduate, Universitas Muhammadiyah Gresik, 2017. Accessed: Dec. 02, 2022. [Online]. Available: <http://digilib.umg.ac.id/gdl.php?mod=browse&op=read&id=jipptumg--lolitaapri-2745&q=LOLITA%20APRILLA>
- [16] N. nasrul, "Analisis dan Perancangan Sistem Informasi Pemetaan Kompetensi Pengajar LP3 STT-NF Berbasis Web Menggunakan Yii Framework", *J. Teknologi Terpadu*, vol. 2, no. 2, 2017.



SISTEM INFORMASI PELAYANAN SURAT MENYURAT DI KELURAHAN DESA SRIWIJAYA LAMPUNG TENGAH

Andronias Siregar¹, Arief Satriansyah², Rachmat Hidayat³, Maya Septa Wijaya⁴

¹Teknologi Komputer, Universitas Bina Sarana Informatika

^{2,4}Sistem Informasi, Universitas Bina Sarana Informatika

³Teknologi Informatika, Universitas Bina Sarana Informatika
Jakarta Pusat, DKI Jakarta, Indonesia 12140

andronias.aoe@bsi.ac.id, arief.aie@bsi.ac.id, rachmat.rch@bsi.ac.id, maia.septawijaya0409@gmail.com

Abstract

Information technology has progressed quite rapidly, and more and more years are increasingly developing technology worldwide. Information technology greatly benefits the world government by creating village websites and building computer-based information systems. With information technology, everyone can process and access data and information they want to convey to the whole community through the website so that everyone can access it easily without having to come and ask the staff or person concerned. The problem with the mailing service in the Village of Sriwijaya Village, Central Lampung, where the process of making letters is still done manually, there are still writing errors. The purpose of creating this correspondence information system is to overcome existing problems and help employees and staff make it easier to update village and community development information by easily accessing this information; with this system, staff can also improve skills in information technology. The method used by the waterfall in making this information system includes making correspondence and information systems in the Sriwijaya village, using the PHP program and MYSQL database. The results of this study are a web-based correspondence information system that is very helpful for community services in making certificates, KK, KTP, SKU, and more quickly and saves time.

Keywords: Information system, Letter application, Village web, Waterfall, Website

Abstrak

Teknologi informasi sekarang ini sudah memiliki kemajuan yang cukup pesat dan semakin bertambahnya tahun semakin berkembang teknologi di dunia. Teknologi informasi memberi manfaat yang besar dalam dunia pemerintahan, sebagai contoh dalam pembuatan *website* desa dan pembangunan sistem informasi berbasis komputer. Dengan teknologi informasi setiap orang dapat mengolah dan mengakses data dan informasi yang ingin di ingin disampaikan kepada seluruh masyarakat melalui laman *website* sehingga setiap orang dapat mengaksesnya dengan sangat mudah, tanpa harus datang dan bertanya kepada *staff* atau orang yang bersangkutan. Permasalahan pelayanan surat menyurat di Kelurahan Desa Sriwijaya Lampung Tengah, di mana proses pembuatan surat masih dilakukan pencatatan manual, masih terdapat kesalahan tulisan. Tujuan dibuatnya sistem informasi surat menyurat ini adalah untuk mengatasi permasalahan yang ada serta membantu karyawan dan *staff* mempermudah untuk memperbarui informasi perkembangan desa dan masyarakat mudah dalam mengakses informasi tersebut, dengan adanya sistem ini *staff* juga bisa meningkatkan keterampilan dalam bidang teknologi informasi. Metode yang digunakan *waterfall* dalam pembuatan sistem informasi ini, meliputi pembuatan surat menyurat dan sistem informasi yang ada di desa Sriwijaya, dengan menggunakan program PHP dan *database* MYSQL. Hasil dari penelitian ini yaitu, sistem informasi surat menyurat berbasis web yang sangat membantu pelayanan masyarakat dalam pembuatan surat keterangan, KK, KTP, SKU dan lebih cepat serta menghemat waktu.

Kata kunci: Aplikasi surat, Sistem informasi, Web desa, *Waterfall*, *Website*

1. PENDAHULUAN

Di era globalisasi ini teknologi sudah sangat maju di berbagai bidang seperti pendidikan, kesehatan, pemerintahan dan lain-lain. Di bidang pemerintahan sistem informasi sangat dibutuhkan agar setiap orang dapat dengan

mudah mengakses data ataupun informasi yang ada salah satunya data dan informasi desa. Layanan publik bernilai tambah lebih dikenal dengan teknologi informasi dan komunikasi atau *e-government*[1].

E-Government adalah area aplikasi yang mencakup banyak area relevansi praktis langsung, seperti: Implementasi sistem teknologi informasi, *realisasi* keuntungan, keamanan informasi, kesenjangan *digital*, *akuntabilitas*, *interoperabilitas*[2].

Pelayanan publik adalah suatu kegiatan atau serangkaian kegiatan yang berkaitan dengan pemenuhan kebutuhan pelayanan hukum dan peraturan dari semua warga negara dan penduduk atas barang, jasa, atau layanan pemerintah yang disediakan oleh penyedia layanan publik[3].

Wana Jaya adalah sebuah desa di Kabupaten Subang, Jawa Barat. Administrasi kependudukan harus datang langsung ke kantor desa. Permasalahannya adalah respons yang lambat, kurangnya informasi yang akurat dari desa ke warga, sistem format *excel*, manual *sheet*, proses pengambilan data yang lama dan kurangnya *update* [4].

Pelayanan surat menyurat di Desa kelurahan Sriwijaya lampung tengah mencakup meliputi Surat keterangan kartu tanda penduduk (KT), Kartu Keluarga (KK), Surat Keterangan Tidak Mampu (SKTM), Surat Keterangan Perusahaan (SKU). Pelayanan Desa Sriwijaya masih dilakukan secara manual menggunakan Microsoft *word* serta pengarsipannya masih menggunakan buku sehingga membutuhkan waktu yang cukup lama untuk proses pencarian data yang tersimpan. Staf di kelurahan Desa Sriwijaya masih kurang hanya beberapa orang saja, proses pelayanan juga dilakukan oleh satu orang staf, yang mengakibatkan pelayanan membutuhkan waktu yang lama.

Tujuan penelitian untuk membantu dan memudahkan *staff* kelurahan untuk surat menyurat dengan menggunakan aplikasi web, diharapkan dengan dibuatkan aplikasi membantu masyarakat dan staf dalam pekerjaan, penulis membuat sistem informasi ini menggunakan bahasa pemrograman PHP dan *database* MSQYL dengan metode *waterfall*.

2. METODE PENELITIAN

2.1 Metode pengembangan sistem

Penulis menggunakan metode *waterfall* pengembangan sistem metode *Waterfall* terbagi menjadi empat tahapan yaitu[5] :

1. *Kebutuhan Sistem (System Requirements)*
Selama *fase* ini, penulis melakukan wawancara, diskusi, atau penelitian langsung untuk mendapatkan informasi tentang kebutuhan, ada beberapa macam *requirement* (kebutuhan) menurut *Sommerville* yaitu Kebutuhan pengguna, sistem dan perangkat lunak[6].
2. *Implementation*
Penulis terlebih dahulu membuat program kecil yang disebut unit. Ini akan diintegrasikan ke dalam unit selanjutnya, dikembangkan dan diuji fungsionalitasnya. Ini disebut pengujian unit.

3. *Verification*

Pada *fase* ini, kami menguji apakah sistem sepenuhnya memenuhi persyaratan dan kemudian memvalidasi sistem.

4. *Operation & Maintenance*

Tahap akhir sistem yang sudah jadi dan dijalankan dilakukan pemeliharaan. Pemeriksaan dan perbaikan kesalahan yang tidak ditemukan pada aplikasi yang berjalan.

2.2 Metode pengembangan sistem

Penulis menggunakan dua metode dalam melakukan pengumpulan data adalah :

1. *Observasi*
Peneliti melakukan kegiatan observasi dengan mengamati dan mempelajari secara langsung segala aspek yang berkaitan dengan pengabdian kepada masyarakat di kantor desa Sriwijaya provinsi Lampung Tengah.
2. *Wawancara*
Penulis melakukan proses wawancara terhadap salah satu *staff* di Kantor Kelurahan Desa Sriwijaya mengenai informasi mengenai beberapa data desa.

2.3 Landasan Teori

1. *Sistem*
Terdiri dari unsur-unsur yang saling terintegrasi dengan tujuan yang sama untuk mencapai tujuan yang diharapkan [7].
2. *Sistem Informasi*
Kombinasi terorganisir dari orang, perangkat keras, perangkat lunak, jaringan komunikasi, dan sumber daya data yang mengumpulkan, mengubah, dan mendistribusikan informasi dalam suatu organisasi[8].
3. *Waterfall*
Model *Waterfall* bersifat sistematis dan berurutan dalam membangun sebuah perangkat lunak. Proses pembuatannya mengikuti alur dari mulai analisis, desain, kode, pengujian dan pemeliharaan[9].
4. *Surat menyurat*
Surat merupakan salah satu alat komunikasi tertulis sehari-hari di dalam sebuah organisasi. Jenis surat yang diterima maupun yang dikirim (keluar) berbagai macam, sehingga arus surat dengan frekuensi tinggi sering menimbulkan kemacetan dan kekacauan pada arus informasi. Untuk itu dibutuhkan pengelolaan surat dengan teknik, metode dan sistem tertentu sesuai instansi masing-masing [10].
5. *Website*
Kumpulan halaman yang menampilkan informasi data tekstual, data gambar diam atau bergerak, data animasi, suara, video, dan setiap kombinasi statis dan dinamis,

masing-masing membentuk rangkaian bangunan yang saling berhubungan, masing-masing merupakan jaringan halaman (*hyperlink*)[11].

6. Web Server

Web server yaitu komputer yang digunakan untuk menyimpan dokumen-dokumen web, komputer ini akan melayani permintaan dokumen web dari kliennya. *Web browser* seperti *explorer* atau *navigator* akan berkomunikasi melalui jaringan dengan *web server* Dengan menggunakan HTTP, *browser* akan meminta sebuah layanan yang disediakan oleh *server*. *Server* akan memberikan layanannya jika tersedia juga dengan menggunakan protokol HTTP[12].

7. Entity Relationship Diagram (ERD)

Deskripsi data yang dimodelkan dalam diagram dan digunakan untuk mendokumentasikan data dengan menentukan isi setiap entitas dan hubungan antar entitas[11].

8. Unified Model Language (UML)

Gambar atau grafik untuk memvisualisasikan, mendokumentasikan, membangun, dan menentukan sistem perangkat lunak yang dikembangkan dengan cara berorientasi objek [13]. UML memiliki standar untuk pemetaan sistem cetak biru, yang terdiri dari deskripsi proses yang memetakan kelas ke program yang lebih konkret, basis data tertanam, dan komponen lain yang diperlukan oleh sistem.

9. Personal Home Page (PHP)

PHP berasal dari kata “*Hypertext Preprocessor*”, yaitu “Bahasa pemrograman universal untuk penanganan pembuatan dan pengembangan sebuah situs web dan dapat digunakan secara bersamaan dengan HTML[14].

10. MYSQL

MySQL adalah sistem manajemen basis data SQL bersifat *open source*. MySQL dapat didefinisikan sebagai sistem manajemen *database*. *Database* sendiri merupakan struktur penyimpanan data. Untuk menambah, mengakses dan memproses data yang disimpan dalam sebuah *database* komputer, diperlukan sistem manajemen *database* seperti MySQL *server*[15].

2.4 Analisa Kebutuhan Sistem

Analisis kebutuhan sistem untuk pengguna :

1. Kebutuhan sistem *staff* kelurahan

Untuk mengakses sistem diminta melakukan *login* untuk masuk ke menu utama untuk mengelola surat pengajuan untuk KTP, KK, SKTM, SKU, SKP. *Staff* kelurahan mengecek data masyarakat. *Staff* kelurahan dapat melakukan verifikasi data jika sudah benar, *staff* kelurahan kemudian membuat surat pernyataan dan mencetak surat pernyataan tersebut, membuat laporan serta memberikan informasi data masyarakat.

2. Kebutuhan sistem masyarakat

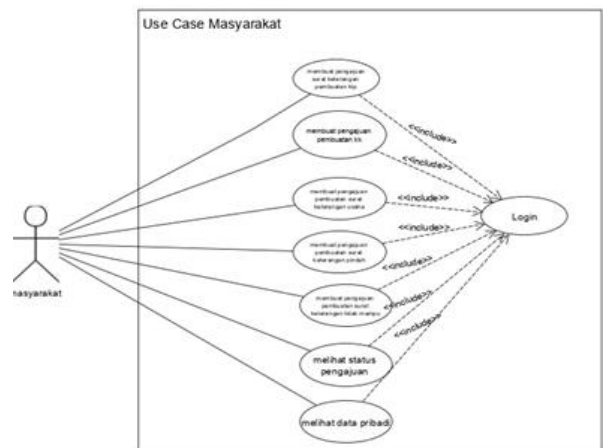
Untuk mengakses ke sistem di minta *login* untuk masuk ke menu utama, setelah itu pengguna dapat mengakses *form* pengajuan surat keterangan KTP, KK, keterangan usaha, keterangan pindahan, keterangan tidak mampu, dan masyarakat dapat melihat informasi status data yang diajukan.

3. HASIL DAN PEMBAHASAN

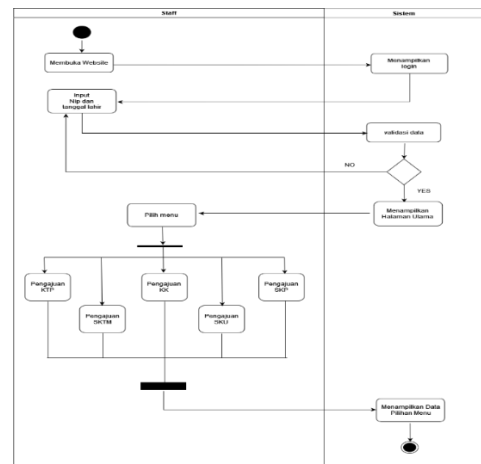
3.1 Use case dan activite diagram Penggunaan sistem untuk masyarakat.

Keterangan pada gambar 1 dan 2

1. Pengguna harus melakukan *login* untuk dapat masuk ke halaman utama.
2. Setelah berhasil pengguna dapat memilih menu pengajuan KTP, KK, SKU, SKP, SKTM, Sistem menampilkan *Form* data pengajuan.
3. Pengguna mengisi *Form* data pengajuan dan mengirim data kemudian sistem akan menyimpan dan data akan di verifikasi oleh *staff*.
4. Sistem akan mengirim informasi persetujuan pengajuan KTP, KK, SKU, SKP, SKTM.



Gambar 1. Use Diagram Penggunaan Sistem

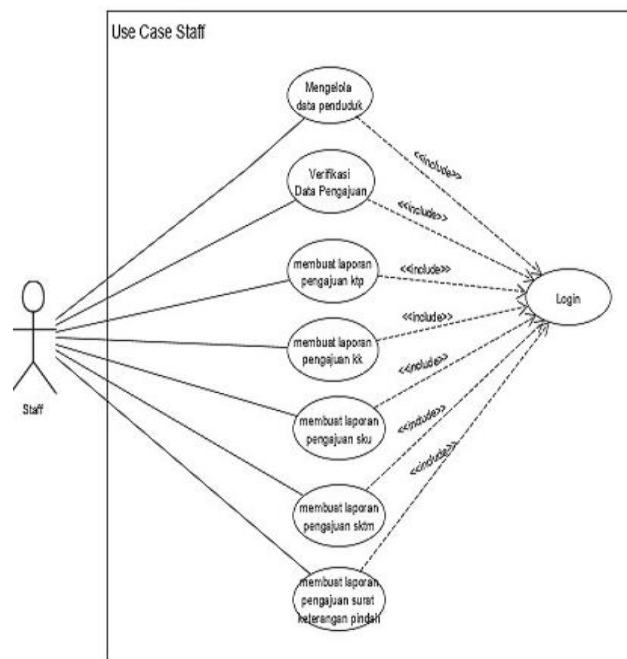


Gambar 2. Activity Diagram Penggunaan Sistem

3.2 Use case dan activite diagram Penggunaan sistem untuk staf kelurahan.

Keterangan pada gambar 3

1. Staf kelurahan harus melakukan *login* untuk dapat masuk ke halaman utama
2. Sistem menampilkan sub menu
3. *Staff* memilih sub menu verifikasi data
4. Sistem menampilkan daftar pengajuan berdasarkan sub menu yang dipilih
5. *Staff* memilih data pengajuan yang ingin diverifikasi
6. Sistem menampilkan data yang dipilih
7. *Staff* melakukan pengecekan dan menyesuaikan data
8. *Staff* menyetujui data yang sudah benar.

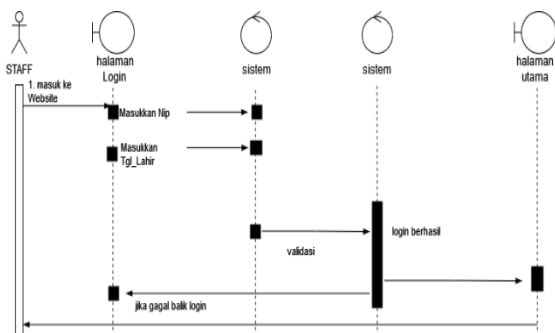


Gambar 3. Use Case Diagram Staf Kelurahan

3.3 Sequence Diagram

1. Sequence Diagram Login

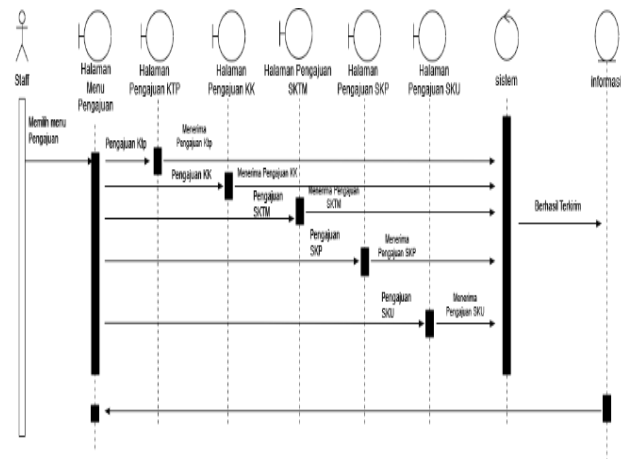
Sequence Diagram Login dapat dilihat pada gambar 4.



Gambar 4. Sequence Diagram Login

2. Sequence Diagram Kelola Surat Pengajuan

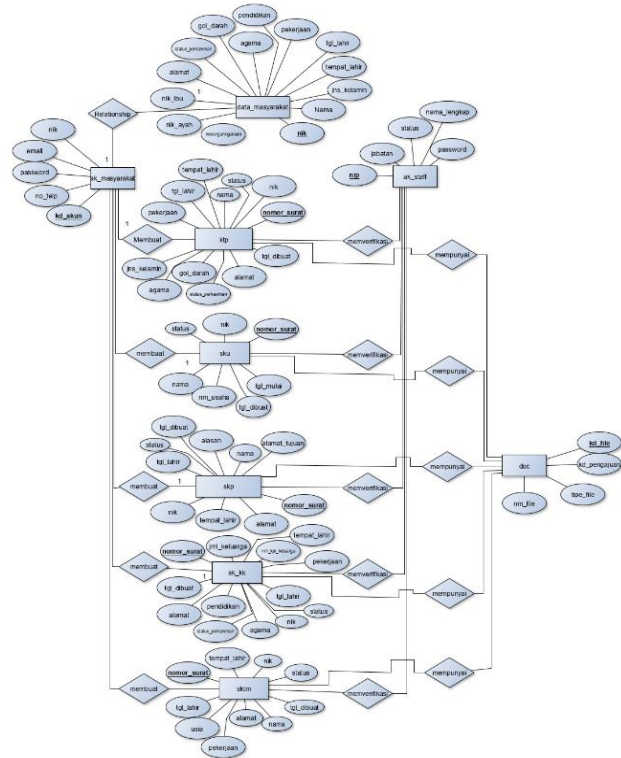
Sequence Diagram Kelola Surat Pengajuan dapat dilihat pada gambar 5 berikut.



Gambar 5. Sequence Diagram Kelola Surat Pengajuan

3.4 Entity Relationship Diagram (ERD)

ERD menggambarkan sebuah relasi antara entitas yang satu dengan yang lain. ERD dapat dilihat pada gambar 6.



Gambar 6. Entity Relationship Diagram

3.5 Rancangan program usulan

1. Halaman Utama

Pada gambar 7 terdapat halaman utama yang merupakan halaman yang ditampilkan saat pengguna pertama kali masuk website, pada halaman pengguna terdapat menu login untuk staff dan masyarakat, serta terdapat menu daftar untuk masyarakat.



Gambar 7. halaman utama

2. Halaman Daftar

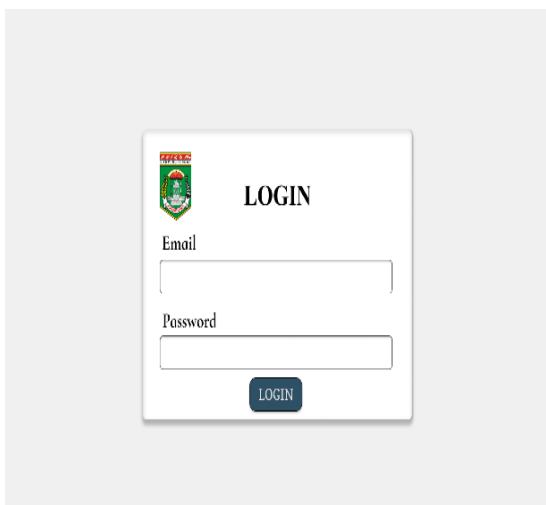
Pada gambar 8, akan tampil halaman pendaftaran bagi pengguna yang belum memiliki akun, pengguna diminta melengkapi data email, nomor telepon, NIK, Password.



Gambar 8. halaman daftar akun

3. Halaman Login

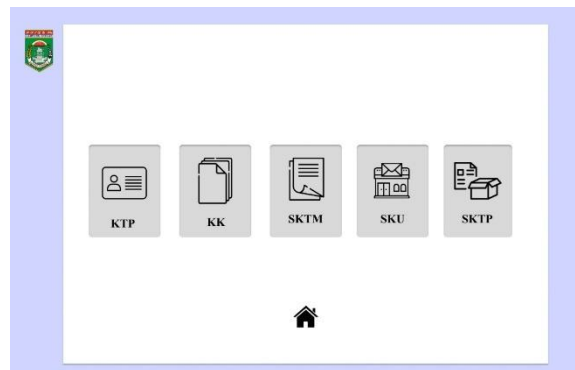
Pada gambar 9 menu login, pengguna diminta memasukkan email dan password untuk masuk ke halaman utama



Gambar 9. Halaman Login

4. Halaman Pelayanan Pengajuan Surat untuk masyarakat

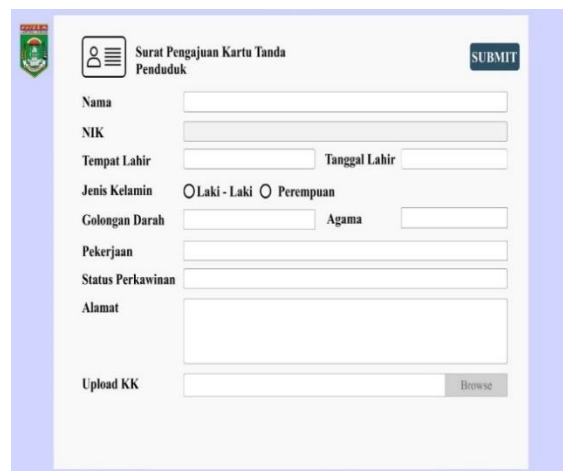
Pada gambar 10, setelah berhasil login akan tampil halaman pengajuan surat untuk masyarakat, pengguna dapat mengajukan pelayanan surat keterangan KTP, KK, SKTM, SKU, SKTP



Gambar 10. Halaman Login

5. Halaman Pengisian surat keterangan KTP

Gambar 11 pengguna diminta mengisi data-data pengajuan surat Keterangan KTP



Gambar 11. Halaman pengisian data surat keterangan KTP

6. Halaman informasi pengisian data

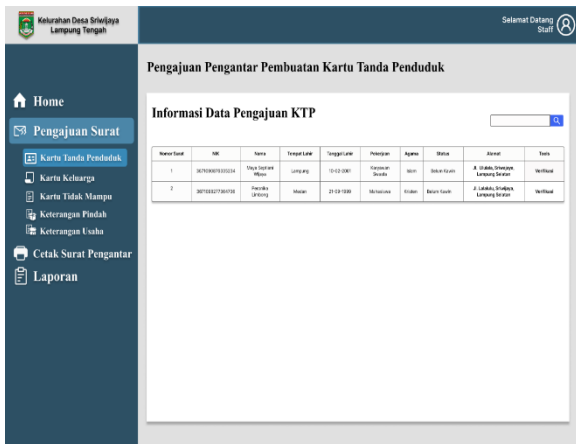
Gambar 12 menampilkan data informasi pengajuan yang telah di isi oleh masyarakat, pengguna dapat melihat status pengajuannya.



Gambar 12. Halaman pengisian data surat keterangan KTP

7. Halaman informasi data-data pengajuan surat

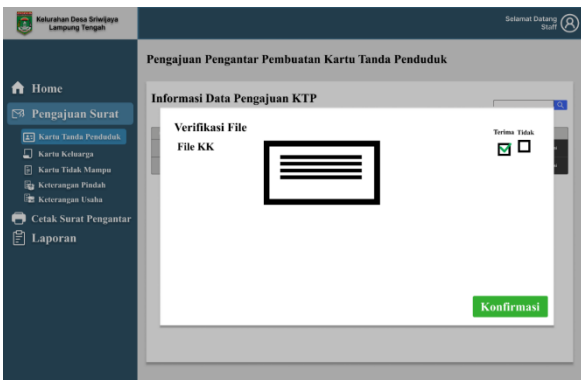
Gambar 13 *staff* melihat informasi data-data pengajuan surat dari masyarakat dan melakukan proses verifikasi.



Gambar 13. Halaman pengisian data surat keterangan KTP

8. Halaman Verifikasi pengajuan oleh staff

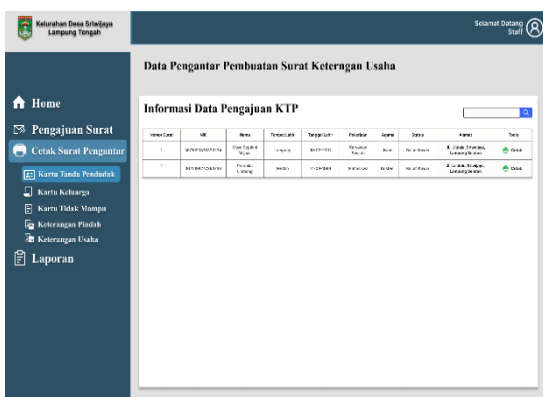
Gambar 14 *staff* akan memverifikasi pengajuan dari masyarakat, jika data pengajuan sesuai dan lengkap *staff* akan menerima dan akan menolak jika tidak lengkap.



Gambar 14. Halaman verifikasi pengajuan surat keterangan

9. Tampilan Halaman Cetak Surat Pengantar

Gambar 15 *staff* kelurahan mencetak surat pengajuan yang telah di terima.



Gambar 15. Halaman Cetak Surat Pengantar

4. KESIMPULAN

Penulis mengambil kesimpulan berdasarkan penelitian ini yaitu :

1. Dengan adanya aplikasi sistem Informasi Pelayanan surat menyurat menjadi lebih cepat dan akurat.
2. Proses pengumpulan data kependudukan relatif sama, sehingga kami menyarankan untuk menggunakan sistem ini di desa lain.
3. Adanya sistem informasi surat menyurat membuka akses pelayanan publik lebih transparan.
4. Data tersimpan dengan aman dalam *database* dan dapat dicadangkan suatu saat diperlukan.
5. Sistem informasi pelayanan surat menyurat menghemat waktu, meningkatkan kinerja, memproses data yang tersimpan dengan baik dan mempercepat pengambilan data yang dibutuhkan.

Ucapan Terima Kasih

Ucapan terima kasih kepada seluruh staf kelurahan Desa Sriwijaya Kabupaten Lampung Tengah.

DAFTAR PUSTAKA

- [1] R. Abdillah, Ari;Lailiah, Badariatul;Sa’adah, Rabiatus; Pebrianto, “Sistem Informasi Pelayanan Masyarakat Secara Online Pada Kelurahan Sungai Raya,” *Inf. Manag. Educ. Prof.*, vol. 6, p. 21, 2021.
- [2] E. B. Pratama and A. Hendini, “Pemodelan Sistem Informasi Layanan Masyarakat (Silam) Pada Kantor Desa Untuk Meningkatkan Pelayanan,” *Klik - Kumpul. J. Ilmu Komput.*, vol. 6, no. 1, p. 49, 2019.
- [3] Pemerintah, “Undang-undang (UU) tentang Pelayanan Publik,” Indonesia, 2009.
- [4] I. Widiastuti, “Sistem Informasi Pelayanan Desa Berbasis Web di Desa Wanajaya Jawa Barat,” *Pendidik. Masy. dan Pengabd.*, vol. 3, p. 887, 2022.
- [5] A. A. Wahid, “Analisis Metode Waterfall Untuk Pengembangan Sistem Informasi,” *J. Ilmu-ilmu Inform. dan Manaj. STMIK*, no. November, pp. 1–5, 2020.
- [6] I. Sommerville, *Software engineering*, 9th ed. 2011.
- [7] N. A. Rahmawati and A. C. Bachtiar, “Analisis dan perancangan sistem informasi perpustakaan sekolah berdasarkan kebutuhan sistem,” *Berk. Ilmu Perpust. dan Inf.*, vol. 14, no. 1, p. 76, 2018.
- [8] E. Anggraeni Yunaeti, *Pengantar Sistem Informasi*, 1st ed. Yogyakarta: andi, 2017.
- [9] R. Pressman, *Rekayasa Perangkat Lunak, Pendekatan*. Yogyakarta: Andi Offset, 2015.
- [10] I. Dewi, Chrisyanti, *Manajemen Kesekretariatan, Printing 1*. Jawa Timur: Prestasi Pustaka, 2019.

- [11] R. L. Khasanah, C. Kesuma, and R. Wijianto, "Sistem Informasi Pelayanan Kesehatan Online Berbasis Web Pada PMI Kabupaten Purbalingga," *Evolusi J. Sains dan Manaj.*, vol. 6, no. 2, pp. 74–83, 2018.
- [12] A. B. Adiwisastro, M. F., & Hikmah, *Web Programming Desain Halaman Web dengan CSS*. Graha Ilmu., 2020.
- [13] H. Nuryansyah and E. Hermawan, "Perancangan Sistem Informasi Manajemen Ekstrakurikuler Berbasis Web Pada Sekolah Menengah Kejuruan Negeri 5 Kota Bandung," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 10, no. 3, pp. 298–305, 2021.
- [14] M. MF, *Buku Sakti Pemrograman Web Seri PHP*. Yogyakarta: Start Up, 2018.
- [15] R. Widyastuti, W. Indrarti, M. Novaliza, and Rani, "Rancang Bangun Sistem Informasi Inventory Boneka Berbasis Web Studi Kasus Di Toko Istana Boneka Cihampelas Bandung," *Prosisko*, vol. 7, no. 2, pp. 96–101, 2020.



PROGRESSIVE WEB APPS : PENGEMBANGAN DAN STUDI PENERIMAAN PADA MAHASISWA INDONESIA MENGGUNAKAN SCRUM DAN UTAUT

Herman¹, Frederick²

^{1,2} Sistem Informasi, Universitas Internasional Batam
Batam, Kepulauan Riau, Indonesia 29444

herman@uib.ac.id, 1931163.frederick@uib.edu

Abstract

Currently, the usage of mobile applications and the total activity on mobile phones through web browsing or native applications is very high. Most users use native mobile apps to browse the content of specific industries. Another way to do so is through a web browser. However, both have limitations. In web applications, users' experience is not that great compared to native applications. On mobile native applications, it requires higher development costs to ensure the app can be accessed across multiple platforms. To solve this problem, Google launched progressive web apps as an alternative, where progressive web apps can be accessed through different platforms, saving development costs but still providing a user experience almost the same as native apps. The focus of this research is to develop a progressive web app with React.js using the scrum method and also to research user acceptance of progressive web apps through qualitative methods/interviews based on the UTAUT research model to general students and lecturers of computer science. The result of this research, based on UTAUT points, shows positive responses toward accepting progressive web apps. Then with the rapid development of the web and browsers, progressive web apps will potentially have a promising future. With this research, it is hoped that can bring benefits to the decision to use the progressive web apps, become a reference for the next research, and increase the knowledge of the public and readers about the progressive web apps.

Keywords: mobile application, progressive web app, react.js, scrum, UTAUT

Abstrak

Saat ini, penggunaan aplikasi *mobile* dan total aktivitas pada ponsel melalui penjelajahan web atau melalui aplikasi *native* sangat tinggi. Sebagian besar pengguna menggunakan aplikasi *native mobile* untuk menelusuri isi industri tertentu. Cara lain adalah dengan melalui *browser* web, namun keduanya memiliki keterbatasan. Pada penjelajahan web, *user experience* yang diterima pengguna tidak terlalu bagus dibandingkan dengan yang diterima pada *native app*, dan pada aplikasi *native mobile*, memerlukan biaya pengembangan yang lebih tinggi untuk memastikan aplikasi bisa diakses pada berbagai platform. Untuk mengatasi masalah tersebut, Google meluncurkan *progressive web app* sebagai sebuah alternatif, di mana *progressive web apps* bisa diakses di berbagai platform yang berbeda, menghemat biaya pengembangan namun tetap bisa memberi *user experience* yang hampir sama dengan *native app*. Fokus dari penelitian ini adalah mengembangkan sebuah *progressive web apps* dengan React.js menggunakan metode *scrum*, serta meneliti penerimaan pengguna terhadap *progressive web apps* melalui metode kualitatif/wawancara dengan model UTAUT kepada mahasiswa dan dosen ilmu komputer secara umum. Hasil penelitian yang berdasar atas poin-poin UTAUT, menunjukkan respons yang positif terhadap penerimaan *progressive web apps*. Kemudian dengan pengembangan-pengembangan cepat dari web dan *browser*, *progressive web apps* berpotensi besar memiliki masa depan yang menjanjikan. Dengan adanya penelitian ini, diharapkan bisa membawa manfaat untuk menilai penggunaan *progressive web apps*, menjadi acuan penelitian berikutnya, dan meningkatkan wawasan masyarakat dan pembaca mengenai *progressive web apps*.

Kata kunci: aplikasi *mobile*, *progressive web app*, react.js, *scrum*, UTAUT

1. PENDAHULUAN

Saat ini, sekitar 60% pengguna di seluruh dunia menggunakan *smartphone*, dan total aktivitas pada ponsel melalui penjelajahan web atau melalui aplikasi *native*

hampir mendekati 67% yang diidentifikasi oleh survei terkini. Sebagian besar pengguna menggunakan aplikasi *native mobile* untuk menelusuri isi industri tertentu. Cara lain adalah dengan melalui *browser* web. Namun keduanya

memiliki keterbatasan. Pada penjelajahan web, tampilan sebagian besar situs web tidak otomatis responsif di web *browser* saat mengakses menggunakan ponsel menyebabkan pengguna tidak merasa senang dan memberi pengalaman yang memuaskan bagi pengguna [1]. Pada aplikasi *native mobile*, dengan pesatnya perkembangan teknologi, membuat perusahaan perlu untuk memastikan aplikasi yang dikembangkan bisa diakses pada perangkat apa pun baik itu di ponsel, komputer, tablet, ataupun perangkat lainnya, terlepas dari sistem operasinya, ukuran layarnya ataupun kondisi perangkatnya. Oleh karena itu memerlukan biaya pengembangan yang lebih mahal dan tinggi. Untuk mengatasi masalah keterbatasan ini, Google meluncurkan *progressive web app* sebagai sebuah alternatif, di mana *progressive web app* bisa diakses di berbagai perangkat terlepas dari sistem operasi, ukuran layar dan kondisi perangkatnya, tanpa perlu mengembangkan beberapa aplikasi secara spesifik di platform yang berbeda. Bahkan untuk *developer* yang mengembangkan aplikasi hanya untuk *iOS* dan *Android*, ini mengartikan penghematan waktu dan tenaga yang besar dalam pengembangan suatu aplikasi. Selain itu, *progressive web app* menawarkan *user experience* yang sangat mirip dengan aplikasi *native* [2].

Progressive Web App (PWA) adalah sebuah teknologi baru yang dirancang dan dikembangkan oleh Google pada Juni 2015 untuk mengatasi keterbatasan *browser* seluler dan aplikasi *native*. *Progressive web app* menggunakan kemampuan web modern yang menggambarkan koleksi teknologi, konsep desain, dan API web yang bekerja bersama-sama untuk menghadirkan pengalaman pengguna seperti aplikasi *native* [3]. *Progressive web app* bisa diakses oleh *mobile users* karena dapat diakses melalui *home screen* dan bekerja terlepas dari koneksi internet. *Progressive web app* juga menawarkan fitur seperti akses kamera, *data storage*, GPS, sensor gerakan dan *push notifications*, yang mana mirip dengan pengalaman yang dialami dari *native apps*. Terlebih lagi, sebuah penelitian menunjukkan bahwa *progressive web app* membutuhkan lebih sedikit

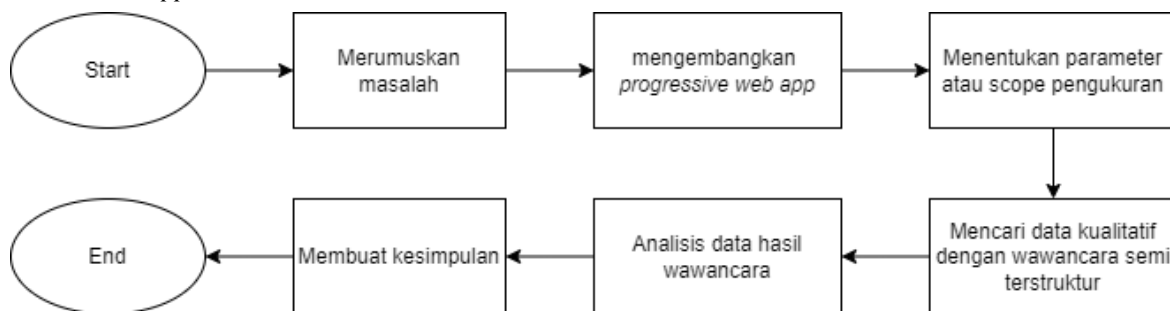
penyimpanan memori dan memiliki waktu peluncuran lebih cepat dibanding *hybrid apps*. Dikarenakan uniknya fitur dan pengalaman seperti *native apps*, *progressive web app* dipercaya bisa memberi *user experience* yang sama baiknya dengan *native apps* [4].

Scrum adalah kerangka kerja ringan yang memungkinkan kemampuan tim untuk mengatasi tantangan adaptif yang kompleks untuk pengembangan dan penyampaian dari produk dengan nilai-nilai yang tinggi dengan meningkatkan kolaborasi, kreativitas dan produktivitas. *Scrum* ditemukan oleh Jeff Sutherland dan Ken Schwaber pada tahun 1993 [5]. Mengetahui bahwa industri teknologi menunjukkan ketertarikan dan berinvestasi pada *progressive web app* untuk mempelajari kemajuan dan peningkatan di masa depan, ditemukan kekurangan penelitian akademik mengenai bagaimana penerimaan pengguna terhadap *progressive web app* [6].

Berdasarkan latar belakang di atas, penelitian ini berfokus pada pengembangan sebuah *progressive web app* dengan React.js menggunakan metode *scrum*, serta memberikan hasil penelitian tentang penerimaan pengguna terhadap *progressive web app* baik pada pengguna dari ponsel ataupun pengguna dari *desktop* melalui metode wawancara kepada mahasiswa dan dosen ilmu komputer secara umum.

2. METODE PENELITIAN

Penelitian dimulai dengan merumuskan masalah berdasarkan topik yang diangkat untuk judul penelitian ini, kemudian mengembangkan sebuah *progressive web app* menggunakan teknologi React.js dengan metode *scrum*. Lalu menentukan parameter atau *scope* pengukuran, yang dilanjutkan dengan merancang pertanyaan-pertanyaan wawancara atau instrumen kualitatif berdasarkan model penelitian UTAUT, kemudian hasil dari wawancara yang didapat, akan melalui proses analisis dari penulis. dan di akhir penulis akan membuat kesimpulan dari penelitian ini. Berikut adalah sebuah ilustrasi garis besar dari alur penelitian ini pada Gambar 1.



Gambar 1. Alur Penelitian

Dalam pengembangan *progressive web app* menggunakan teknologi React.js dan metode *scrum*, pada penelitian ini *development team* akan diperankan oleh penulis. Kemudian untuk *product owner* & *scrum master* akan diperankan oleh dosen pembimbing dari penulis. Berikut adalah detail proses *scrum* pada penelitian ini :

Berikut terlampir *product backlog* dibawah ini

- a. *Setup Development Environments dan Sourcecode*
- b. *Build UI + Responsive UI Developments*
- c. *Build Functionalities*
 - 1) *Search Functionality + Reset Search Keywords*

- 2) Integrate Weather API & Get Details for Every Fetch
- 3) Error Handlings
- 4) Display Background Based on Weather Condition Result
- d. PWA Developments + Lighthouse Testing
- e. Testings & Bug Fixes
- f. App Deployment

Pada penelitian ini, sebuah *sprint* akan berdurasi 2 minggu. Rencana dari perancangan aplikasi adalah dengan 2 *sprint* yang memiliki tugas yang terbagi sesuai di bawah ini

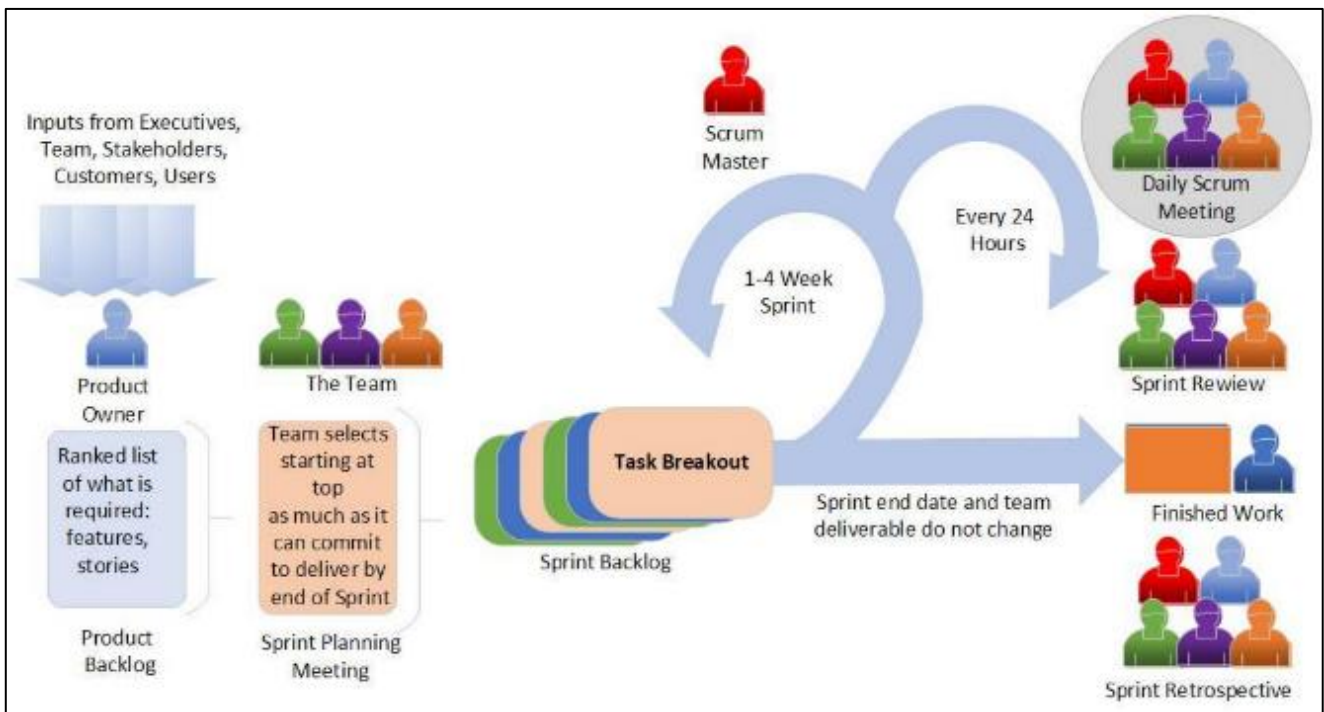
Sprint Backlog 1

- a. Setup Development Environments dan Sourcecode
- b. Build UI + Responsive UI Developments
- c. Build Functionalities

Sprint Backlog 2

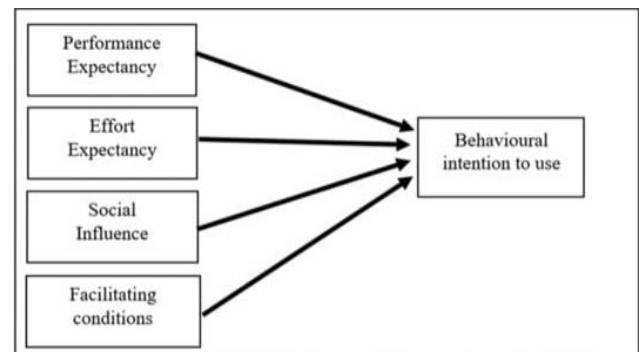
- a. PWA Developments + Lighthouse Testing
- b. Testings & Bug Fixes
- c. App Deployment

Dalam menjalankan proses/tahap *sprint* dalam *scrum*, penulis juga melakukan *daily scrum* bersama *product owner* & *scrum master* penulis, kemudian setelah berakhirnya masa waktu sebuah *sprint* / 2 minggu, penulis melakukan *sprint review* dan *sprint retrospective* bersama dosen pembimbing. Sebab tidak ada fitur tambahan dan proses pengembangan berjalan lancar, maka *sprint* tetap berlanjut seperti biasa. Hingga di akhir *sprint* ke-2, proses pengembangan menghasilkan sebuah *progressive web app* yang telah dikerahkan ke *third-party provider*, yaitu *Netlify* dan bisa diakses oleh masyarakat secara *online*. Berikut adalah contoh ilustrasi *scrum* menurut [7] pada Gambar 2 :



Gambar 2. Ilustrasi Scrum

Setelah menyelesaikan pengembangan *progressive web app*, penulis melakukan pengumpulan data menggunakan metode kualitatif atau wawancara dengan konsep semi terstruktur yang berdasar atas model UTAUT kepada dosen/mahasiswa ilmu komputer secara umum, dengan jumlah sebesar 30 atas referensi dari [8]. UTAUT memiliki 4 kunci konstruksi model yaitu: harapan kinerja (*performance expectancy*), harapan usaha (*effort expectancy*), pengaruh sosial (*social influence*), dan kondisi fasilitas (*facilitating conditions*) yang memiliki pengaruh terhadap niat perilaku untuk menggunakan teknologi [9]. Berikut adalah ilustrasi model UTAUT pada Gambar 3.



Gambar 3. Model UTAUT

Dengan UTAUT dan 4 kunci konstruksi modelnya, penulis menyusun daftar pertanyaan wawancara berdasarkan referensi dari [10], [11] dan [12] sebagai berikut.

Tabel 1. Daftar Pertanyaan Wawancara

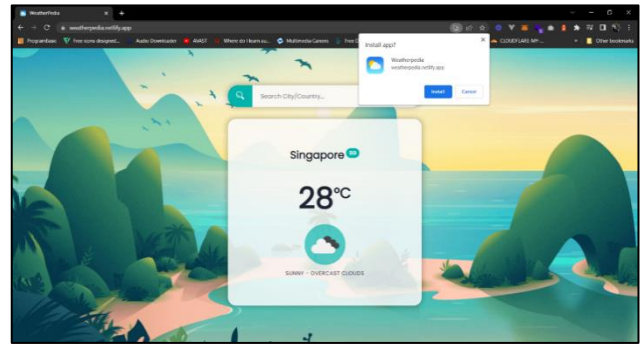
Variabel	Pertanyaan Wawancara
Ekspektasi Kinerja	1. Apakah pengguna merasakan pengalaman yang sama seperti saat menggunakan aplikasi <i>native</i> lainnya saat mengakses <i>progressive web app</i> ini? 2. Apakah pengguna merasa nyaman / menerima <i>progressive web app</i> ini? 3. Apakah <i>progressive web app</i> memberikan pilihan yang fleksibel kepada pengguna untuk mengakses aplikasi ini dari website ataupun langsung dari <i>home screen</i> ?
Ekspektasi Usaha	1. Apakah menggunakan <i>progressive web app</i> terasa lebih praktis jika dibandingkan dengan <i>native app</i> yang diinstall dari Playstore ataupun <i>install</i> langsung ke <i>desktop</i> ? 2. Apakah pengguna ada merasakan ketidaknyamanan atau penolakan terhadap <i>progressive web app</i> ini?
Pengaruh Sosial	1. Seberapa setuju pengguna terhadap pernyataan bahwa orang yang memengaruhi perilaku pengguna atau orang terdekat, berpikir/menyarankan bahwa sebaiknya pengguna mengakses segala aplikasi secara langsung dari <i>home screen</i> tanpa menggunakan <i>website</i> , layaknya <i>native apps</i> / <i>desktop apps</i> ?
Kondisi yang Memfasilitasi	1. Seberapa lancar <i>device</i> pengguna dalam menggunakan <i>progressive web app</i> ini? 2. Apakah pengguna mengetahui bahwa <i>website</i> yang adalah <i>progressive web app</i> , mempunyai tombol unduh/download, yang bila ditekan, bisa menambahkan <i>progressive web app</i> tersebut ke dalam <i>home screen</i> ?
Keinginan untuk mencoba <i>progressive web app</i>	1. Seberapa berminat Anda untuk mencoba dan beraktivitas di dalam <i>progressive web app</i> lainnya yang bisa diakses dari <i>website</i> maupun langsung dari <i>home screen</i> , yang cocok dengan kebutuhan Anda?

Hasil wawancara dari Tabel 1 akan dianalisis dengan menggunakan eksploratif deskriptif dengan cara menyampaikan hasil wawancara secara umum.

3. HASIL DAN PEMBAHASAN

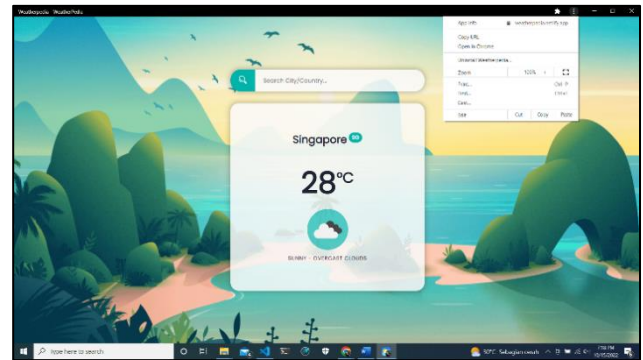
Berikut adalah beberapa cuplikan hasil pengembangan *progressive web app* menggunakan React.js dengan metode *scrum* :

Pada Gambar 4, menunjukkan tampilan *progressive web app* yang sedang diakses melalui *browser* Google Chrome di *desktop*, yang memiliki pilihan *install* ke *desktop*.



Gambar 4. Tampilan di *Desktop*

Pada Gambar 5, menunjukkan hasil/tampilan *progressive web app* yang telah diunduh kedalam *desktop* dan digunakan layaknya *desktop app* lainnya.

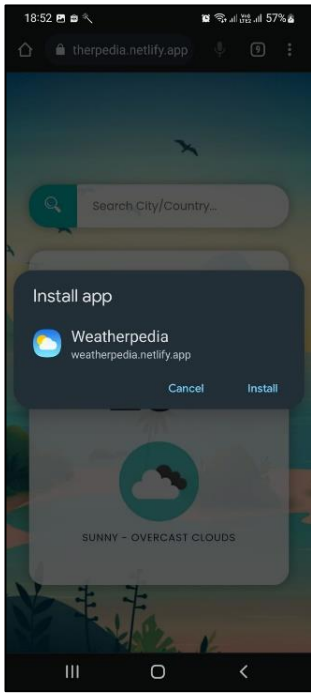


Gambar 5. Tampilan di *Desktop*

Pada Gambar 6 dan 7, menunjukkan tampilan *progressive web app* yang diakses melalui *browser* Google Chrome pada ponsel yang memiliki pilihan untuk *install* ke *mobile*.

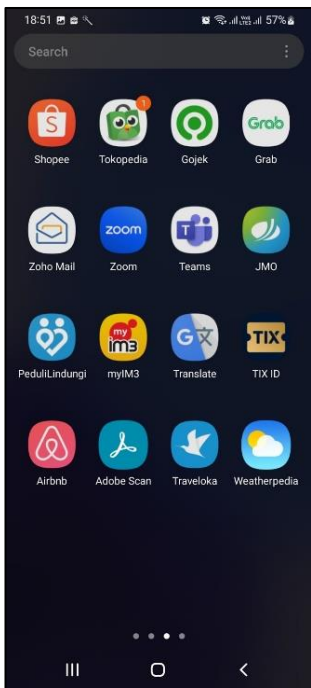


Gambar 6. Tampilan *Website* dibuka di Ponsel



Gambar 7. Tampilan Opsi Instalasi pada Ponsel

Pada Gambar 8 dan 9, menunjukkan tampilan *progressive web app* yang telah diinstall ke *mobile* dan hasil akses *progressive web app* tersebut pada *mobile*.



Gambar 8. Tampilan Aplikasi Terinstal di Ponsel



Gambar 9. Tampilan Aplikasi Terinstal Dibuka di Ponsel

Hasil wawancara kepada 31 orang terkait *progressive web app* direpresentasikan menggunakan *affinity diagram*. *Affinity diagram* dipilih karena memungkinkan untuk mengelompokkan pola berulang dalam data, menambah struktur dan memperoleh kejelasan masalah secara efisien. Selain itu, teknik *affinity diagram* juga direkomendasikan dalam jurnal ini untuk menganalisis data yang telah dikumpulkan selama wawancara, *focus group discussions*, maupun studi lapangan [4]. Berikut adalah *affinity diagram* yang telah dikategorikan berdasar atas poin-poin UTAUT.

Tabel 2. *Affinity Diagram* Ekspektasi Kinerja

Ekspektasi Kinerja
"...menurut saya aplikasinya sudah sangat mirip dengan aplikasi <i>native</i> lainnya."
"...saya pernah menggunakan <i>progressive web app</i> , menurut saya di <i>progressive web app</i> ini terasa nyaman-nyaman saja dan praktis."
"...bagi saya <i>progressive web app</i> sangat fleksibel, karena saya bisa mengakses dari <i>website/home screen</i> . Bila saya butuh cepat untuk mengakses <i>progressive web app</i> , <i>progressive web app</i> ini sangat berguna sebab bisa kita tambahkan ke <i>home screen</i> ."

Dari segi Ekspektasi Kinerja pada Tabel 2, rata-rata orang menjawab bahwa pengalaman yang dirasakan dari *progressive web app* terasa sama dengan *native apps*, mereka merasa nyaman dalam menggunakan *progressive web app*, dan juga merasa bahwa *progressive web app* fleksibel karena bisa diakses dari *website* maupun *home screen*.

Tabel 3. Affinity Diagram Ekspektasi Usaha

Ekspektasi Usaha
"... <i>progressive web app</i> terasa praktis bagi saya, <i>overall</i> juga mirip dengan <i>native apps</i> ."
"...iya! Saya merasa <i>progressive web app</i> praktis, juga dari <i>progressive web app</i> yang kita saksikan ini juga sangat enak dilihat, tapi memang lebih enak aja gitu, karena bisa diakses dari <i>home screen</i> ataupun dari <i>browser</i> ."
"...bagi saya tidak ada penolakan/ketidaknyamanan terhadap <i>progressive web app</i> ."

Dari segi Ekspektasi Usaha pada Tabel 3, rata-rata orang menjawab bahwa *progressive web app* praktis bagi mereka, dan tidak ada ketidaknyamanan ataupun penolakan terhadap *progressive web app*.

Tabel 4. Affinity Diagram Pengaruh Sosial

Pengaruh Sosial
"...kalau menurut saya ada plus minusnya sih, kalau dari untuk <i>home screen</i> nya kan kita perlu <i>download</i> makan <i>space</i> , kalau dari <i>website</i> perlu buka akses internet. Jadi netral."
"...kalau saya lebih prefer <i>home screen</i> karena kalau <i>website</i> kan kita perlu buka <i>browser</i> dan isi url atau semacamnya, jadi saya lebih prefer <i>home screen</i> sih."
"...saya kurang setuju bila mengakses segala aplikasi langsung dari <i>home screen</i> tanpa menggunakan <i>website</i> sebab tidak semua <i>device</i> punya <i>storage</i> yang cukup, jadi pakai <i>website</i> bisa membantu."

Dari segi Pengaruh Sosial pada Tabel 4, rata-rata orang menjawab bahwa mereka setuju dengan pendapat untuk mengakses kebanyakan aplikasi dari *home screen*, karena faktor seperti kenyamanan, kemudahan, dll. Terdapat juga beberapa orang yang menjawab bahwa mereka kurang setuju/netral terhadap pendapat tersebut.

Tabel 5. Affinity Diagram Kondisi yang Memfasilitasi

Kondisi yang Memfasilitasi
"... <i>device</i> saya sangat lancar dalam menggunakan <i>progressive web app</i> ini."
"...cukup lancar karena aplikasinya juga ringan."
"...saya mengetahui lokasi untuk tombol <i>download</i> pada <i>progressive web app</i> , biasanya di Google Chrome di <i>handphone</i> terletak di bawah dengan kata-kata seperti <i>add to home screen</i> ."

Dari segi Kondisi yang Memfasilitasi pada Tabel 5, rata-rata orang menjawab bahwa mengakses *progressive web app* terasa sangat lancar bagi mereka, dan tidak ada kendala, serta mereka mengetahui lokasi dan cara *download progressive web app* ke *home screen* mereka.

Tabel 6. Affinity Diagram Keinginan untuk Mencoba PWA

Keinginan untuk mencoba menggunakan <i>progressive web app</i>
"...kalau saya sama dengan teman-teman, cukup tertarik, cukup berminat juga."
"...ngelihat <i>weather app</i> yang menggunakan <i>progressive web app</i> , dan terlihat lebih praktis, jadi pasti sangat minat sih untuk melihat <i>progressive web app</i> lainnya."
"...minat atau tidaknya mencoba <i>progressive web app</i> lainnya tergantung kebutuhan saya, bila aplikasinya cocok dengan kebutuhan saya tentu saya akan mencoba."

Dari segi Keinginan untuk Mencoba menggunakan *Progressive Web App* pada Tabel 6, rata-rata orang menjawab bahwa mereka tertarik dan berminat untuk mencoba dan mengakses *progressive web app* lainnya.

4. KESIMPULAN

Pada penelitian ini, *Scrum & React.js* terbukti dapat digunakan untuk mengembangkan *progressive web app*. Dengan selesainya pengembangan sebuah *progressive web app*, dilakukan pengumpulan data menggunakan metode kualitatif atau wawancara. Hasil penelitian yang diperoleh telah direpresentasikan dalam bentuk *affinity diagram* yang telah dikategorikan berdasar atas poin-poin UTAUT. Secara tren dan suara mayoritas, hasil penelitian menunjukkan respon yang positif terhadap penerimaan *progressive web app*. Dengan pengembangan-pengembangan cepat dari web dan *browser*, *progressive web app* akan memiliki masa depan yang menjanjikan untuk berpotensi menggantikan *native apps* dan menjadi kebutuhan dasar situs web apa pun di masa depan [13]. Dengan demikian penerimaan terhadap *progressive web app* yang telah menunjukkan tren positif, berpotensi untuk berkembang lebih pesat di masa depan [14].

Sebagai pekerjaan berikutnya atau di masa depan, akan bagus bila bisa menganalisis perspektif yang lebih beragam dari kelompok pengguna yang lebih luas (contohnya pengguna anak-anak, pengguna orang dewasa, pengguna orang lanjut usia, dll). Penelitian lebih lanjut mengenai keamanan dalam menggunakan *progressive web app* juga akan menjadi penelitian yang membantu penelitian pada *progressive web app* menjadi lebih lengkap [15].

Ucapan Terima Kasih

Berikut adalah ucapan terima kasih penulis kepada semua pihak yang telah membantu dan memberikan kontribusi dalam menyelesaikan penelitian ini.

Tuhan yang Maha Esa sebagai kepercayaan penulis, atas rahmat dan karunia-Nya, penulis dapat menyelesaikan penelitian ini.

Kedua orang tua dari penulis, yang senantiasa memberi dukungan kepada penulis.

Bapak Dr. Iskandar Itan, S.E., MM., sebagai Rektor, Universitas Internasional Batam.

Bapak Dr Hendi Sama sebagai Dekan Fakultas Ilmu Komputer, Universitas Internasional Batam.

Bapak Tony Wibowo, S.Kom., MMSI, sebagai ketua program studi sistem informasi, Universitas Internasional Batam

Bapak Herman, S.Kom., M.Kom sebagai dosen pembimbing yang membimbing dan meluangkan waktu untuk membimbing penulis dalam pembuatan penelitian ini.

Teman-teman dari penulis yang tidak bisa disebutkan satu per satu.

DAFTAR PUSTAKA

- [1] A. Mhaske, A. Bhattad, P. Khamkar, and R. More, "Progressive Web App for Educational System," *Int. Res. J. Eng. Technol.*, pp. 310–312, 2018, [Online]. Available: <https://arc.applause.com/2015/11/30/application-shell->.
- [2] I. Rajesh N and S. John Jacob, "IRJET-Progressive Web Apps: A lighter alternative Cite this paper Progressive Web Apps: A lighter alternative," *Int. Res. J. Eng. Technol.*, 2020, [Online]. Available: www.irjet.net.
- [3] N. Nurwanto, "Penerapan Progressive Web Application (PWA) pada E-Commerce," *Techno.Com*, vol. 18, no. 3, pp. 227–235, 2019, doi: 10.33633/tc.v18i3.2400.
- [4] J. Sedkowska and D. R. Kiunsi, "How does the user experience of a progressive web application compare to native application?," 2020, [Online]. Available: <https://www.diva-portal.org/smash/get/diva2:1455626/FULLTEXT01.pdf>.
- [5] K. Bhavsar, V. Shah, and S. Gopalan, "Scrum: An Agile Process Reengineering in Software Engineering," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 3, pp. 840–848, 2020, doi: 10.35940/ijitee.c8545.019320.
- [6] V. Sharma, R. Verma, V. Pathak, M. Paliwal, and P. Jain, "Progressive Web App (PWA) - One Stop Solution for All Application Development Across All Platforms," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, pp. 1120–1122, 2019, doi: 10.32628/cseit1952290.
- [7] J. Mihajlović Milićević, F. Filipović, I. Jezdović, T. Naumović, and M. Radenković, "Scrum Agile Framework in E-business Project Management: An Approach to Teaching Scrum," *Eur. Proj. Manag. J.*, vol. 9, no. 1, pp. 52–60, 2019, doi: 10.18485/epmj.2019.9.1.7.
- [8] K. Vasileiou, J. Barnett, S. Thorpe, and T. Young, "Characterising and justifying sample size sufficiency in interview-based studies: Systematic analysis of qualitative health research over a 15-year period," *BMC Med. Res. Methodol.*, vol. 18, no. 1, pp. 1–18, 2018, doi: 10.1186/s12874-018-0594-7.
- [9] S. Rahi and M. Abd.Ghani, "Investigating the role of UTAUT and e-service quality in internet banking adoption setting," *TQM J.*, vol. 31, no. 3, pp. 491–506, 2019, doi: 10.1108/TQM-02-2018-0018.
- [10] T. W. Ningrum, "Pengaruh Ekspektasi Kinerja, Ekspektasi Usaha, Faktor Sosial, Kesesuaian Tugas dan Kondisi yang Memfasilitasi Pemakai terhadap Minat Pemanfaatan Sistem Informasi," *Simp. Nas. Sist. Teknol. Inf.*, pp. 1–23, 2016.
- [11] M. A. Almaiah, M. M. Alamri, and W. Al-Rahmi, "Applying the UTAUT Model to Explain the Students' Acceptance of Mobile Learning System in Higher Education," *IEEE Access*, vol. 7, pp. 174673–174686, 2019, doi: 10.1109/ACCESS.2019.2957206.
- [12] W. Khoirunnisak, "Implementasi Model Penerimaan Unified Theory Of Acceptance And User Of Technology (UTAUT) Untuk Menganalisis Faktor-Faktor Penerimaan Dosen Terhadap Penggunaan E-Learning Share-ITS," *Tesis*, pp. 200–202, 2016.
- [13] K. A. Sedek, "Progressive Web," *Wikipedia*, 2019, [Online]. Available: https://en.wikipedia.org/wiki/Progressive_web_applications.
- [14] G. de Andrade Cardieri and L. M. Zaina, "Analyzing user experience in mobile web, native and progressive web applications: A user and HCI specialist perspectives," *ACM Int. Conf. Proceeding Ser.*, 2018, doi: 10.1145/3274192.3274201.
- [15] D.-J. Rensema, "The Current State of Progressive Web Apps," 2019, [Online]. Available: <https://www.browserlondon.com/blog/2019/04/15/current-state-progressive-web-app-pwa/>.



PERBANDINGAN MODEL *MACHINE LEARNING* PADA KLASIFIKASI TUMOR OTAK MENGGUNAKAN FITUR *DISCRETE COSINE TRANSFORM*

Simeon Yuda Prasetyo¹, Ghinaa Zain Nabiilah²

^{1,2} School of Computer Science, Universitas Bina Nusantara
Jakarta Barat, DKI Jakarta, Indonesia 11480

simeon.prasetyo@binus.ac.id, ghinaa.nabiilah@binus.ac.id

Abstract

Brain tumors are abnormal tissue growths characterized by excessive cell growth in certain brain parts. One of the reliable techniques currently available to identify brain tumors is using Magnetic Resonance Imaging (MRI) scans. The scanned MRI images are monitored and examined for tumor detection by a specialist. Developing more effective and efficient tools to help medical professionals identify brain tumors is urgent as the number of people suffering from brain tumors soars, and the death rate will reach 18,600 in 2021. In previous research, machine learning-based models demonstrated the ability to detect brain tumors with a classification accuracy of 92%, and this result is reliable. We computationally tested several hyperparameters using publicly available MRI datasets to obtain the most reliable binary classification accuracy in MRI brain images. A high level of model accuracy is achieved by testing various existing machine-learning model architectures and inserting a feature map extracted from the Discrete Cosine Transform (DCT). Classification of MRI images achieved the highest accuracy on test data at 93% using the Support Vector Machine (SVM) model.

Keywords: Binary Classification, Brain Tumor, DCT, Machine Learning, SVM

Abstrak

Tumor otak adalah pertumbuhan jaringan abnormal yang ditandai dengan pertumbuhan sel yang berlebihan di bagian otak tertentu. Salah satu teknik andal saat ini yang tersedia untuk mengidentifikasi tumor otak adalah penggunaan pemindaian *Magnetic Resonance Imaging* (MRI). Gambar MRI yang dipindai dipantau dan diperiksa untuk deteksi tumor oleh dokter spesialis. Mengembangkan alat yang lebih efektif dan efisien untuk membantu profesional medis mengidentifikasi tumor otak dirasa cukup mendesak karena jumlah orang yang menderita tumor otak melonjak dan tingkat kematian yang mencapai 18.600 pada tahun 2021. Dalam penelitian sebelumnya, model berbasis pembelajaran mesin mampu menunjukkan kemampuan untuk mendeteksi tumor otak dengan akurasi klasifikasi 92% dan hal hasil ini dapat diandalkan. Untuk mendapatkan akurasi klasifikasi biner yang paling andal dalam gambar otak MRI, kami menguji secara komputasi beberapa *hyperparameter* menggunakan kumpulan data MRI yang tersedia untuk umum. Tingkat akurasi model yang tinggi dicapai dengan menguji berbagai arsitektur model *machine learning* yang ada, diikuti dengan memasukkan *feature map* yang diekstraksi dari *Discrete Cosine Transform* (DCT). Klasifikasi gambar MRI mencapai akurasi pada data *test* tertinggi sebesar 93% dengan menggunakan model *Support Vector Machine* (SVM).

Kata kunci: DCT, Klasifikasi Biner, *Machine Learning*, SVM, Tumor Otak

1. PENDAHULUAN

Tumor adalah sel yang tumbuh secara tidak normal (proliferasi sel berlebihan) di bagian tubuh tertentu. Tumor adalah jaringan abnormal yang dapat berbentuk padat ataupun cair[1].

Menurut SEER (2021), kasus kanker dan tumor otak diperkirakan 24.530 kasus baru pada tahun 2021. Kematian akibat kanker otak mencapai 18.600. Selain itu, angka

harapan hidup bagi mereka yang terkena penyakit ini adalah 32,6% dalam 5 tahun[2].

Mengidentifikasi adanya tumor otak bukanlah hal yang mudah karena gejala yang dialami oleh penderita tumor otak dapat menyebabkan gangguan kejiwaan seperti kecemasan dan depresi [3]. Oleh karena itu, gangguan kejiwaan biasanya terdiagnosis ketika pasien memiliki tumor otak. Secara umum dalam diagnostik tumor otak, sebagian dari tumor otak pertama kali dideteksi menggunakan *magnetic*

resonance imaging (MRI) yang merupakan teknologi andal saat ini. Gambar MRI yang dipindai dipantau dan diperiksa oleh ahli radiologi dan profesional medis atau dokter spesialis untuk menentukan apakah ada tumor. Oleh karena itu, menentukan apakah seorang pasien mengidap tumor otak dapat memakan waktu lama dan memerlukan analisis menyeluruh oleh tim profesional medis yang dapat membaca dan mencitrakan tumor dari pencitraan MRI. Selain itu, kompleksitas gambar MRI dapat menyebabkan penentuan negatif palsu saat mendeteksi keberadaan tumor otak. Oleh karena itu, untuk mengatasi kelemahan yang signifikan ini, pengembangan perangkat medis menggunakan pembelajaran mesin untuk tidak hanya mempercepat seluruh proses identifikasi, tetapi juga untuk membuat keputusan yang lebih tepat tentang mendeteksi adanya tumor otak.

Penelitian sebelumnya [4] menggunakan teknik pengolahan citra biner dan morfologi citra. Klasifikasi tumor diterapkan setelah segmentasi dilakukan dengan proses ekstraksi ciri bentuk. Hasil klasifikasi tumor yang dicapai adalah 89,5%, yang dapat memberikan informasi yang lebih jelas dan spesifik untuk deteksi tumor. Hasil ini diperoleh dengan menggunakan klasifikasi *K-Nearest Neighbor*. Studi lain dalam proses klasifikasi [5] menganalisis sifat statistik dari gambar masukan, mengklasifikasikan data secara sistematis ke dalam kategori yang berbeda, dan kemudian membandingkan tiga algoritma klasifikasi: JST, *random forest*, dan SVM. Menggunakan SVM memberikan akurasi terbaik yakni 90%.

Di sisi lain, percobaan [6] membandingkan dua algoritma yang berbeda untuk fase segmentasi: segmentasi *U-network* dan segmentasi berbasis CNN. Di sisi lain, tumor *grading* membandingkan algoritma VGG-16 dan CNN. Berdasarkan percobaan yang dilakukan untuk hasil terbaik menggunakan segmentasi U-Net, algoritma deteksi tumor menggunakan *random forest classifier*, dan tumor *grading level* menggunakan algoritma VGG-16. Eksperimen ini menggunakan spektroskop MR dan *dataset* FLAIR berhasil mencapai akurasi eksperimen tertinggi sebesar 92%.

Studi lain mengklasifikasikan citra otak MRI menjadi tiga kelas tumor: glioma, meningioma, dan hipofisis. Algoritma untuk mengekstrak peta fitur tingkat tinggi dari model berbasis VGG16 menggunakan Fast R-CNN sebagai pengklasifikasi. Studi ini mencapai akurasi rata-rata 77,6% di semua kelas [7]

Ayesha Samreen, dkk [8] juga melakukan penelitian terkait dengan klasifikasi citra otak MRI. *Dataset* yang digunakan senada dengan *dataset* pada penelitian ini yaitu berasal dari *kaggle* dan mengklasifikasikan citra menjadi dua jenis, yaitu tumor otak dan normal. Akurasi optimalnya adalah menggunakan CNN dengan hasil akurasi 0.9290.

Selain itu, penggunaan CNN untuk mengklasifikasi citra otak MRI, juga dilakukan oleh Nyoman Abiwinanda, dkk [9]. Penelitian ini menggunakan CNN untuk mengenali tiga

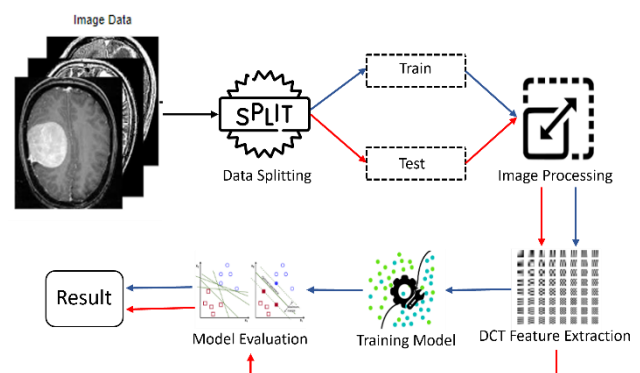
jenis tumor otak, yaitu Glioma, Meningioma, dan Pituitari. Hasil optimal penelitian ini adalah dengan nilai akurasi 84.19% pada data validasi.

Ekstraksi fitur DCT digunakan dalam penelitian ini karena masih sedikit penelitian yang memanfaatkan fitur DCT ini, sehingga masih perlu dilakukan eksplorasi lebih lanjut, terutama pada tugas klasifikasi tumor otak. DCT mencari koefisien yang memiliki kemampuan lebih untuk membedakan kelas yang berbeda secara lebih baik daripada koefisien lainnya. Selain itu, DCT juga memiliki rentang koefisien yang luas [10].

Penelitian ini bertujuan untuk mengembangkan alat diagnostik tumor otak berbasis citra MRI menggunakan ekstraksi ciri bentuk *Discrete Cosine Transform* (DCT) yang merupakan algoritma kompresi sinyal, dan melakukan pengujian beberapa model *machine learning* pada beberapa ukuran citra untuk mendapatkan model klasifikasi yang terbaik. Hasil akhir dari penelitian ini adalah penyajian matriks konfusi dari model terbaik yang dibangun dan hasil berupa persentase akurasi pengujian.

2. METODE PENELITIAN

Gambar 1 menyajikan Tahapan penelitian yang digunakan dalam penelitian ini.



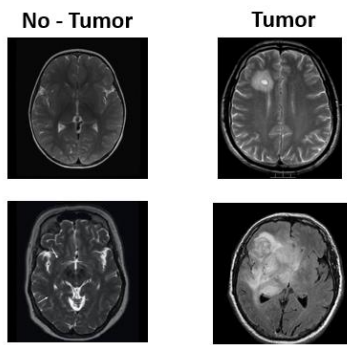
Gambar 1. Tahapan Penelitian

Penjelasan dari proses Tahapan Penelitian dari Gambar 1 yakni:

2.1 Pengumpulan Data

Pada tahap pengumpulan data ini dilakukan pencarian *dataset* yang relevan terhadap penelitian ini. *Dataset* yang dikumpulkan merupakan data citra MRI tumor otak yang terdapat label berupa adanya tumor ataupun tidak ada tumor.

Pada tahapan ini menggunakan *dataset* yang berisikan citra MRI tumor otak yang dapat diakses melalui situs *kaggle* [11] dengan *dataset* bernama *BraTS* 2019. Pada *dataset* ini terdapat citra dengan jumlah total 3000 citra MRI otak, di mana 1500 citra merupakan citra MRI dengan tumor dan 1500 citra lainnya merupakan citra MRI yang tidak terdapat tumor. Beberapa contoh citra dari *dataset BraTS* 2019 dapat dilihat pada Gambar 2.



Gambar 2. Contoh Dataset BraTS 2019

Gambar 2 berisi contoh *dataset* yang berisi citra MRI yang termasuk tumor dan bukan tumor. Karakteristik dari data yang termasuk tumor terdapat area lesi sebagai ilustrasi dari tumor, seperti terlihat pada Gambar 2 bagian kanan. Sedangkan untuk kondisi normal yang bukan tumor terlihat pada Gambar 2 bagian kiri.

2.2 Pembagian Data

Data dilakukan pembagian menjadi 2 bagian secara merata untuk proses *training* dan testing atau pengujian. Data dibagi menjadi data *training* sejumlah 2800 data (1400 citra MRI dengan tumor, dan 1400 Citra MRI tanpa tumor) dan 200 data untuk testing (100 merupakan citra MRI dengan tumor, dan 100 merupakan citra MRI otak tanpa tumor).

2.3 Pra-Pemrosesan Data

Pada tahapan ini citra diubah menjadi aras keabuan (*single channel color*) yang mana semula citra memiliki 3 *channel* RGB. Selanjutnya, pada tahapan pra pemrosesan ini juga melakukan perubahan citra yang semula memiliki ukuran yang tidak sama menjadi berukuran 150x150 piksel ataupun 224 x 224 *pixel* (ukuran piksel ini nantinya juga akan dibandingkan).

2.4 Ekstraksi Fitur DCT

Pada klasifikasi citra memerlukan pemrosesan ekstraksi fitur untuk dapat mengubah citra yang semula memiliki fitur sangat banyak dan tidak memiliki signifikansi yang tinggi (memiliki korelasi yang rendah terhadap *output*) menjadi fitur yang lebih ringkas namun berarti atau tingkat signifikansi yang lebih tinggi. Dalam penelitian ini, DCT (*Discrete Cosine Transform*) digunakan sebagai ekstraksi fitur dari citra MRI tumor otak yang ada.

Algoritma DCT (*Discrete Cosine Transform*) merupakan algoritma yang biasa digunakan dalam kompresi sinyal (gambar merupakan salah satu bentuk dari sinyal digital). Algoritma ini melakukan perubahan data spasial menjadi bentuk frekuensi, kemudian dilakukan pengolahan data frekuensi yang kemudian dikonversi ke bentuk spasial kembali dengan inversi metode tersebut. Di mana, domain frekuensi merupakan sinyal yang memiliki wujud berupa amplitudo terhadap waktu (spektrum terhadap suatu sinyal waktu), domain frekuensi merujuk pada sinyal yang

berhubungan dengan frekuensi dalam suatu waktu. Sedangkan, domain spasial merujuk pada dasar pada manipulasi piksel secara langsung pada citra [12] [13]

Discrete cosine transform dirumuskan seperti pada persamaan (1).

$$v(k) = \alpha(k) \sum_{n=0}^{N-1} U(n) \cdot \text{Cos} \left[\frac{\pi(2n+1)k}{2N} \right] \quad (1)$$

$$\text{Untuk } 0 \leq k \leq N - 1$$

$$a(w) \left\{ \begin{array}{l} \sqrt{\frac{1}{N}} \text{ untuk } w = 0 \\ \sqrt{\frac{2}{N}} \text{ untuk lainnya} \end{array} \right\}$$

Di mana $v(k)$ merupakan hasil ekstraksi fitur DCT, dengan k yang menyatakan indeks baris atau kolom hasil ekstraksi. N merupakan ukuran baris atau kolom blok citra yang hendak dilakukan ekstraksi fitur DCT. Sedangkan $U(n)$ menyatakan blok citra dengan aras keabuan.

Pada ekstraksi fitur DCT biasanya dapat dilakukan dengan sejumlah tahapan. Tahap pertama adalah mendapatkan citra atau gambar sebagai matriks warna citra (biasanya, perhitungan matriks pada satu komponen warna baik *red* atau *green* atau *blue* ataupun citra keabuan $[red+green+blue] / 3$). Selanjutnya, citra dibagi ke dalam blok berukuran 8 x 8 piksel, urut dari ujung kiri ke kanan, dan dari atas ke bawah.

2.5 Pembentukan Model Machine Learning

Dalam penelitian ini terdapat beberapa model yang akan digunakan untuk diuji dan dipilih berdasarkan akurasi *test* terbaiknya. Beberapa model klasifikasi yang akan digunakan antara lain :

- *Naive Bayes*

Naive Bayes merupakan salah satu algoritma yang biasa digunakan untuk tugas klasifikasi, baik klasifikasi citra ataupun teks. *Naive Bayes* menggunakan pendekatan *probabilistic* yang menggunakan teorema probabilitas *bayes*. Algoritma *Naive Bayes* membuat asumsi kemunculan fitur yang tidak saling bergantung. Sehingga cocok digunakan untuk mendeteksi tumor otak dari berbagai lokasi berbeda dengan fitur berbeda [14]. Penelitian ini menggunakan *Gaussian Naive Bayes*, pada persamaan 2 merupakan bentuk rumus dari *Gaussian Naive Bayes*.

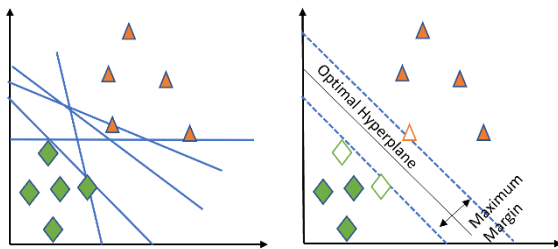
$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma^2y}} \exp \left(-\frac{(x_i-\mu_y)^2}{2\sigma^2y} \right) \quad (2)$$

- *Decision Tree*

Decision Tree adalah salah satu metode yang biasanya digunakan dalam berbagai bidang, seperti klasifikasi citra, pengenalan pola, regresi, dan

klasifikasi teks. Algoritma *Decision Tree* bekerja dengan menggunakan beberapa aturan untuk membuat keputusan yang terstruktur seperti pohon [15].

- *K-Nearest Neighbor* (KNN)
KNN merupakan algoritma klasifikasi yang bekerja dengan menggunakan sejumlah nilai k tetangga atau data terdekat [16]. Perhitungan jarak tetangga terdekat dapat dihitung berdasarkan rumus *Euclidean distance*, *Manhattan distance*, atau *Minkowsky distance*. Algoritma ini bekerja berdasarkan kedekatan atau kemiripan terhadap data lainnya. Jumlah K pada KNN bisa ditentukan sesuai kebutuhan. Pada penelitian ini jumlah K yang digunakan adalah 1, 3, dan 5.
- *Support Vector Machine* (SVM)
SVM adalah salah satu algoritma *supervised learning* yang digunakan untuk klasifikasi ataupun regresi. Cara kerja SVM adalah dengan mencari *hyperplane* atau fungsi pemisah antar kelas terbaik dan memaksimalkan jarak antar kelas [17]. SVM digunakan untuk memaksimalkan margin dalam mencari *hyperplane* yang optimal. Gambar 3 berisi ilustrasi untuk mencari *hyperplane* optimal.



Gambar 3. Ilustrasi Optimal *Hyperplane* SVM

2.6 Evaluasi Model

Pada tahapan evaluasi model, dilakukan perhitungan menggunakan *Confusion Matrix* untuk memperoleh nilai Akurasi, Presisi, *Recall*, dan *F1 Score*. Perhitungan dalam *Confusion Matrix* memperhatikan nilai *true positive* (TP), *true negative* (TN), *false positive* (FP) dan *false negative* (FN).

Akurasi menggambarkan persentase seluruh data yang diklasifikasikan benar pada kelas positif dan negatif. Sehingga proses perhitungan yang dilakukan adalah dengan membagi seluruh data prediksi yang bernilai benar dengan seluruh data yang tersedia. Sehingga jika dirumuskan akan menjadi persamaan (3).

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FN+FP)} \quad (3)$$

Di mana seluruh nilai pada persamaan tersebut bisa dilihat berdasarkan *confusion matrix* untuk memperoleh nilai pada setiap bagian [18]

Precision menggambarkan seberapa tepat atau akurat model yang dibuat. Dengan mengukur prediksi positif (berapa banyak yang benar-benar positif). Sehingga *precision* dapat dirumuskan dalam persamaan (4).

$$Precision = \frac{(TP)}{(TP+FP)} \quad (4)$$

Recall atau *sensitivity* menggambarkan berapa banyak *Actual Positive* yang diprediksi oleh model dengan pelabelan sebenarnya adalah positif. Sehingga *recall* dapat dirumuskan seperti persamaan (5).

$$Recall = \frac{(TP)}{(TP+FN)} \quad (5)$$

F1-score merupakan metrik yang mempertimbangkan keseimbangan rerata secara harmonik antara *precision* dan *recall*. Pada *f1-score* nilai negatif aktual (*true negative*) diabaikan. Metrik ini menggambarkan kekokohan tugas dalam klasifikasi [18]. *F1 score* dapat dirumuskan seperti persamaan (6).

$$F1 - Score = \frac{(2 \times Precision \times Recall)}{(Precision + Recall)} \quad (6)$$

3 HASIL DAN PEMBAHASAN

Dalam penelitian ini evaluasi kinerja model diuji coba di lingkungan Python Google Colab Pro yang dipercepat oleh Unit Pemrosesan Grafis (GPU) dan RAM 25 GB. Menggunakan ekstraksi fitur DCT yang dilanjutkan dengan klasifikasi menggunakan beberapa model menunjukkan hasil yang cukup baik, bahkan terdapat hasil yang sedikit lebih baik dari penelitian sebelumnya yang menggunakan pendekatan *deep learning*.

Percobaan ini dilakukan penentuan parameter untuk sejumlah algoritma. Pada SVM menggunakan *kernel* RBF dengan derajat 3 dengan perhitungan γ seperti pada persamaan (7).

$$\gamma = \frac{1}{(n \text{ features} * X.var())} \quad (7)$$

Sedangkan pada *Decision Tree* menggunakan *Gini index* untuk penentuan *split* kemudian batas *node* diperluas sampai seluruh *leaf* pada *tree* menjadi murni.

Pada *Naive Bayes* perhitungan probabilitas menggunakan pendekatan distribusi normal (*Gaussian*). Sedangkan pada KNN, jarak yang digunakan merupakan jarak *euclidean* dengan jumlah *neighbor* yang diuji adalah 1, 3 dan 5.

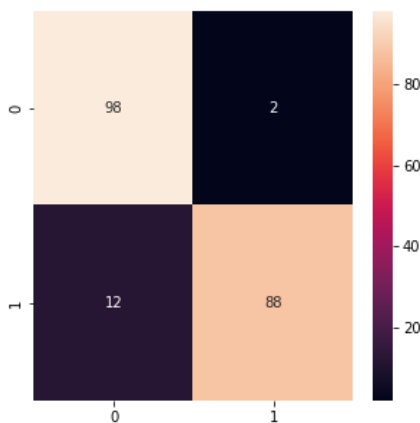
Hasil eksperimen deteksi atau klasifikasi biner tumor otak dengan menggunakan ekstraksi DCT dan klasifikasi *machine learning* dapat dilihat pada tabel 1.

Tabel 1. Hasil Percobaan Model

Model	150 x 150 px		224 x 224 px	
	Training	Testing	Training	Testing
SVM	97.464 %	90.5 %	97.785 %	93.0 %
Tree	100.0 %	79.5 %	100.0 %	83.5 %
Bayes	69.142 %	55.0 %	70.0 %	55.5 %
KNN-1	100.0 %	91.0 %	100.0 %	91.0 %
KNN-3	98.107 %	83.5 %	98.464 %	84.5 %
KNN-5	94.214 %	81.0 %	94.285 %	78.5 %

Berdasarkan hasil pada tabel di atas diketahui bahwa model konfigurasi terbaik diperoleh oleh model SVM dengan ukuran citra 224 x 224 piksel. Pada model tersebut berhasil mendapatkan akurasi *training* sebesar 97.8% dan akurasi *testing* sebesar 93%.

Berdasarkan konfigurasi model terbaik yang dibentuk tersebut mendapatkan *confusion matrix* yang tersaji pada Gambar 4.



Gambar 4. Confusion Matrix

Confusion matrix yang ditampilkan pada Gambar 4, sebelah kiri atas merupakan *True Negative* (TN) sejumlah 98, sebelah kanan atas merupakan *False Positive* (FP) sejumlah 2, sebelah kiri bawah merupakan *False Negative* (FN) sejumlah 12, dan sebelah kanan bawah merupakan *True Positive* (TP) sejumlah 88. Di mana pada sumbu X dan Y terdapat angka nol dan satu, nol merupakan kelas *no-tumor* dan satu merupakan kelas tumor.

Sedangkan hasil perhitungan akurasi berdasarkan nilai *precision*, *recall*, dan *F1 Score* terdapat pada Gambar 5.

	precision	recall	f1-score	support
0	0.89	0.98	0.93	100
1	0.98	0.88	0.93	100
accuracy			0.93	200
macro avg	0.93	0.93	0.93	200
weighted avg	0.93	0.93	0.93	200

Gambar 5. Hasil Perhitungan Evaluasi

Gambar 5 menunjukkan nilai *precision*, *recall*, dan *f1 score* pada metode dengan akurasi *testing* terbaik, yaitu SVM. Dalam gambar tersebut terdapat hasil klasifikasi model pada setiap kelas, yaitu 0 dan 1. Kelas 0 menunjukkan klasifikasi bukan tumor, sedangkan 1 menunjukkan klasifikasi tumor. Sedangkan untuk nilai *F1 Score* dihitung berdasarkan *macro average*. *Macro average* akan menghitung metrik secara independen untuk setiap kelas lalu kemudian mengambil nilai rata-rata. Pada perhitungan *Macro average* semua kelas dihitung secara merata.

4 KESIMPULAN

Telah ditunjukkan bahwa dengan menerapkan ekstraksi fitur menggunakan *Discrete Cosine Transform* (DCT) untuk memperoleh *feature map* dan menggunakan model – model *machine learning* menunjukkan bahwa skor akurasi tertinggi pada data *test* mendapatkan nilai hingga 93%. Akurasi *test* 93% ini didapatkan melalui model SVM (*support vector machine*) dengan ukuran citra sebesar 224 x 224 piksel.

Saran untuk perbaikan model agar ke depannya dapat dioptimalkan dengan mencari *hyperparameter* terbaik dari model tersebut seperti jenis *kernel* yang digunakan dan *kernel degree*. Saran perkembangan selanjutnya, bisa mencoba alternatif model gabungan (konsep *ensemble learning*) yang mungkin bisa meningkatkan performa model.

DAFTAR PUSTAKA

- [1] T. Sinha, "Tumors: Benign and Malignant," *Cancer Ther Oncol Int J*, vol. 10, no. 3, May 2018, doi: 10.19080/ctoj.2018.10.555790.
- [2] SEER, "Surveillance, Epidemiology, and End Results Program," <https://seer.cancer.gov/>.
- [3] S. A. H. S. Javadi and B. Rezaei, "Brain tumors and indications for brain imaging in patients with psychiatric manifestations: a case report," *Middle East Current Psychiatry*, vol. 28, no. 1, Dec. 2021, doi: 10.1186/s43045-021-00136-2.
- [4] Rafi Haidar Ramdlon, Entin Martiana Kusumaningtyas, and Tita Karlita, "Brain Tumor Classification Using Mri Images With K-Nearest Neighbor Method," *2019 International Electronics Symposium (IES)*, pp. 660–667, 2019.

- [5] Göklap Cinarer and Bülent Gürsel Emiroglu, "Classification of Brain Tumors by Machine Learning Algorithms," *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 2019.
- [6] M. A. Naser and M. J. Deen, "Brain tumor segmentation and grading of lower-grade glioma using deep learning in MRI images," *Comput Biol Med*, vol. 121, Jun. 2020, doi: 10.1016/j.compbiomed.2020.103758.
- [7] Yakub Bhanothu, Anandhanarayanan Kamalakannan, and Govindaraj Rajamanickam, "Detection and Classification of Brain Tumor in MRI Images using Deep Convolutional Network," *2020 6th International Conference on Advanced Computing & Communication Systems (ICACCS)*, pp. 248–252, 2020.
- [8] A. Samreen, A. M. Taha, Y. V. Reddy, and P. Sathish, "Brain Tumor Detection by Using Convolution Neural Network," *International journal of online and biomedical engineering*, vol. 16, no. 13, pp. 58–69, 2020, doi: 10.3991/ijoe.v16i13.18545.
- [9] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, "Brain tumor classification using convolutional neural network," in *IFMBE Proceedings*, 2019, vol. 68, no. 1, pp. 183–189. doi: 10.1007/978-981-10-9035-6_33.
- [10] S. Dabbaghchian, A. Aghagolzadeh, and M. S. Moin, "Feature extraction using discrete cosine transform for face recognition," in *2007 9th International Symposium on Signal Processing and its Applications, ISSPA 2007, Proceedings*, 2007. doi: 10.1109/ISSPA.2007.4555358.
- [11] Aryan Felix, "BraTS 2019 (Train/Test/Valid)," 2020.
- [12] P. Verma and G. Verma, "Brain Tumor Detection using Hybrid Model of DCT DWT and Thresholding," *International Research Journal of Engineering and Technology (IRJET), Certified Journal | Page*, vol. 9001, 2008.
- [13] "Discrete Cosine Transform; Second Edition"
- [14] Hein Tun Zaw, Noppadol Maneerat, and Khin Yadanar Win, "Brain tumor detection based on Naïve Bayes Classification," *2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)*, 2019.
- [15] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.
- [16] B. Tu, J. Wang, X. Kang, G. Zhang, X. Ou, and L. Guo, "KNN-Based Representation of Superpixels for Hyperspectral Image Classification," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 11, no. 11, pp. 4032–4047, Nov. 2018, doi: 10.1109/JSTARS.2018.2872969.
- [17] E. García-Gonzalo, Z. Fernández-Muñoz, P. J. G. Nieto, A. B. Sánchez, and M. M. Fernández, "Hard-rock stability analysis for span design in entry-type excavations with learning classifiers," *Materials*, vol. 9, no. 7, Jul. 2016, doi: 10.3390/ma9070531.
- [18] D. M. W. Powers and Ailab, "Evaluation: From Precision, Recall And F-Measure To Roc, Informedness, Markedness & Correlation."



ANALISIS *TEXT MINING* KLASIFIKASI KEGIATAN KELUARGA MENGGUNAKAN ORANGE DENGAN METODE NAÏVE BAYES

Arsya Fathiarahma¹, Apriade Voutama², Taufik Ridwan³, Nono Heryana⁴

^{1,2,3,4} Sistem Informasi, Universitas Singaperbangsa Karawang
Jawa Barat, Karawang, Indonesia 41361

arsya.fathiarahma19005@student.unsika.ac.id, apriade.voutama@staff.unsika.ac.id,
taufik.ridwan@cs.unsika.ac.id, nono@unsika.ac.id

Abstract

The dual roles performed by women often lead to conflict in the family. Disputes can occur due to the lack of family roles in helping balance the work at home and office experienced by career women and mothers. The Naïve Bayes algorithm is used in this study to determine the results of applying the Naïve Bayes algorithm classification based on role, activity, age, and interest in a family. A total of 287 data records generated from the survey were used in this study and a data split of 80:20 for training and testing data using orange data mining tools. The results show that the calculation accuracy with the Naïve Bayes algorithm reaches 93%, with the conclusion that the use of orange data mining tools results in good accuracy.

Keywords: Conflict, Double role, Naïve bayes, Orange data mining tools, Woman

Abstrak

Peran ganda yang dilakukan oleh wanita kerap kali menimbulkan konflik dalam keluarga. Konflik dapat terjadi karena kurangnya peran keluarga dalam membantu keseimbangan antara pekerjaan di rumah dan kantor yang dialami oleh para wanita karir sekaligus ibu. Algoritma *Naïve Bayes* dipergunakan pada penelitian ini dengan tujuan untuk mengetahui hasil penerapan klasifikasi algoritma *Naïve Bayes* berdasarkan peran, aktivitas, umur dan minat dalam sebuah keluarga. Sejumlah 287 *data record* yang dihasilkan dari survei dipergunakan dalam penelitian ini dan dilakukan *split data* 80:20 untuk *data training* dan *data testing* dengan menggunakan *tools orange data mining*. Hasil penelitian menampilkan bahwa akurasi perhitungan dengan algoritma *Naïve Bayes* mencapai 93% dengan kesimpulan bahwa penggunaan *tools orange data mining* menghasilkan akurasi yang baik.

Kata kunci: Konflik, *Naïve bayes*, Peran ganda, *Tools orange data mining*, Wanita

1. PENDAHULUAN

Informasi yang kompleks dan variatif beberapa tahun belakangan ini dihasilkan dari perkembangan sistem informasi[1]. Berbagai peningkatan utilitas sistem informasi disebabkan karena adanya berbagai metode yang ditanamkan pada sebuah *engine*. Salah satu pemanfaatan sistem informasi yaitu ditemukannya informasi baru dari sebuah data dengan digunakannya *text mining*[2]. Terdapat beberapa penggunaan *text mining*, salah satunya yaitu *classifier* dengan *naïve bayes*[3].

Objek pada penelitian ini yaitu berawal dari perubahan pandangan masyarakat akibat tingginya gaya hidup, status sosial, ataupun adanya pergeseran budaya membuat wanita mempunyai peran ganda sebagai ibu rumah tangga sekaligus wanita karir[4]. Hal tersebut dikuatkan oleh jumlah persentase jumlah pekerja perempuan berdasarkan

status perkawinan, jenis kelamin, dan daerah tempat tinggal yang dipaparkan oleh BPS RI tahun 2018[5].

Tabel 1. Persentase Penduduk Usia 15 tahun ke Atas yang Bekerja

	Status Perkawinan				Total
	Belum Kawin	Kawin	Cerai Hidup	Cerai Mati	
Perkotaan					
Perempuan	19,96	66,51	4,39	9,14	100,00
Laki-Laki	23,04	73,33	1,78	1,85	100,00
Perdesaan					
Perempuan	10,17	76,04	3,31	10,48	100,00
Laki-Laki	19,77	76,35	1,65	2,23	100,00
Perkotaan +Perdesaan					

	Status Perkawinan				Total
	Belum Kawin	Kawin	Cerai Hidup	Cerai Mati	
Perempuan	15,39	70,96	3,89	9,76	100,00
Laki-Laki	21,52	74,73	1,72	1,03	100,00

Peran ganda yang dilakukan, akan mempengaruhi kinerja di kantor maupun di rumah yang dapat menimbulkan tekanan pada wanita yang menyebabkan terjadinya konflik atas ketidakseimbangan antara pekerjaan dan keluarga[6]. Permasalahan tersebut juga dapat disebabkan adanya stereotip yang dibuat oleh masyarakat terkait pekerjaan. Stereotip adalah menggeneralisasikan individu ke suatu kelompok tertentu dikarenakan sebuah persepsi atau sebuah penilaian[7].

Kualitas *Feminine* atau maskulin harus dimiliki oleh seseorang agar dianggap sukses dalam melakukan pekerjaan. Karena kualitas *feminine* merupakan kualitas yang diasosiasikan dengan jenis kelamin perempuan, pekerjaan rumah seperti memasak, mencuci baju, membersihkan rumah, kerap kali masyarakat menganggap bahwa perempuan lebih cocok melakukan pekerjaan tersebut[8].

Dalam penelitian ini, *dataset* yang digunakan diperoleh dari hasil survei dengan pertanyaan mengenai minat, keahlian, peran, serta aktivitas atau kegiatan keluarga oleh 80 keluarga di 4 wilayah yang berbeda. Perolehan data didapatkan sebanyak 287 yang akan dibagi menjadi 2, yaitu *data training* dan *data testing* dalam proses pemodelan. Tujuan dalam penelitian ini yaitu menampilkan hasil dari penerapan metode *Naïve Bayes* dalam klasifikasi pada data yang didapatkan dari hasil survei dengan adanya penggunaan sebuah *tools* yaitu *orange data mining*.

Text mining merupakan salah satu bidang perluasan dari *data mining* yang melakukan proses pengidentifikasian pola, topik, atau atribut lainnya pada data yang berbentuk teks semi struktur atau tidak terstruktur dengan mengeksplor dan melakukan analisa yang menghasilkan temuan informasi baru[9]. Secara umum, *text mining* mengacu pada proses ekstrak informasi dari dokumen tidak terstruktur[10]. Pendefinisian *text mining* dapat dijelaskan dengan adanya penemuan informasi baru atau tidak diketahui sebelumnya oleh komputer dan informasi dari berbagai sumber yang berhasil diekstrak merupakan kunci dari proses *text mining*[11].

Salah satu algoritma yang sering digunakan terhadap klasifikasi *text mining* yaitu *Naïve Bayes classifier*. *Teorema bayes* ini merupakan algoritma dengan penggunaan pengalaman sebelumnya untuk memprediksi peluang kedepannya[12]. *Naïve Bayes classifier* mempunyai ciri utama yaitu independensi. Alasan teorema ini disebut naif adalah karena membuat asumsi kemunculan fitur tidak saling bergantung[13]. Bahkan jika terdapat ciri yang bergantung satu sama lain atau muncul atas

keberadaan ciri lainnya, semua sifat ini secara individual berkontribusi pada probabilitas[14]. Secara singkat, algoritma *naïve bayes classification* adalah diprediksinya semua probabilitas untuk pengklasifikasian sekumpulan data algoritma yang sering

digunakan karena dapat memungkinkan dapat bekerja lebih dan kompleks dari keadaan nyata [15]

Rumus *naïve bayes*:

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

Keterangan

X : Belum diketahui *class* pada data

A : Suatu *class* spesifik pada hipotesis data X

$P(Y|X)$: Probabilitas hipotesis berdasar kondisi

$P(Y)$: Probabilitas hipotesis

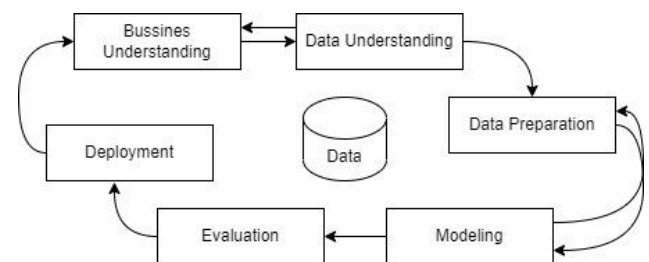
$P(X|Y)$: Probabilitas berdasarkan kondisi pada hipotesis

$P(X)$: Probabilitas Y

Berdasarkan penjelasan tujuan pada penelitian ini, *tools* yang akan digunakan untuk penerapan metode *Naive Bayes* dalam klasifikasi kegiatan keluarga yaitu *orange data mining*. Teknologi *engine open source* ini berguna dalam menganalisis dan menggambarkan data eksploratif. Di dalamnya terdapat banyak pilihan seperti pemodelan prediktif, eksperimen, maupun sistem rekomendasi sebagai bahan penelitian biomedis, genomik, dan atau pengajaran yang membuat *tools* ini mempermudah proses analisa data dengan intuitif[16].

2. METODE PENELITIAN

Kegiatan yang dilakukan sebelum penelitian adalah penyelidikan atau pencarian suatu fakta dengan teliti sesuai metode tertentu agar dapat menjelaskan dan menjawab suatu permasalahan[17]. Pada Gambar 1 merupakan tahap metode penelitian eksperimen, serta metodologi CRISP-DM yang akan dijelaskan sebagai berikut:



Gambar 1. Metode CRISP-DM

2.1. Pemahaman Bisnis (*Business Understanding*)

Pemahaman bisnis yang dimaksudkan pada penelitian ini yaitu pemahaman akan sebab dan akibat dari peran ganda dari seorang wanita yang menjadi seorang ibu, istri, sekaligus wanita karir. Seperti, adanya kesulitan untuk melakukan semua pekerjaannya, pertikaian dalam rumah tangga hanya karena seorang istri telat untuk melayani suaminya, ataupun adanya ketidakseimbangan perlakuan antar anak satu dan lainnya dalam memberikan tanggung jawab pekerjaan rumah. Maka dari itu, strategi awal yang diterapkan adalah melakukan penelitian kepada beberapa keluarga untuk mencari sekiranya apa saja pekerjaan rumah yang dapat dibagi dengan sama rata untuk mengurangi pekerjaan dari seorang wanita berperan ganda maupun pertikaian antar anak

2.2. Pemahaman Data (*Data Understanding*)

Setelah melakukan pemahaman atas permasalahan yang diangkat, selanjutnya melakukan tahapan *data understanding*. Tahapan ini dilakukan dengan menentukan atribut apa saja yang akan dipertanyakan ke keluarga-keluarga dan melakukan survei kepada mereka.

2.3. Persiapan Data (*Data Preparation*)

Tahap pada *data preparation* memiliki *preprocessing* yang akan dilakukan. *Data cleaning* merupakan salah satu *preprocessing* yang dilakukan pada *data preparation* untuk menghapus *missing value* pada salah satu atau beberapa atribut yang ada. Kegunaan dari menghapus *missing value*

pada atribut ialah agar data yang dihitung lebih akurat karena sudah tidak ada *missing value*.

2.4. Pemodelan (*Modeling*)

Setelah data telah dipersiapkan, langkah selanjutnya yaitu tahap pemodelan. Model yang digunakan pada penelitian ini yaitu *Naïve Bayes classifier*. Untuk melakukan model tersebut, dilakukannya pembagian data dengan teknik *split* antara data yang ingin dilatih (*training data*) dan data yang ingin diuji (*testing data*) sejumlah 80:20.

2.5. Evaluasi (*Evaluation*)

Tahap evaluasi yang dimaksud adalah melakukan proses tinjauan dari hasil model yang dilakukan dan perhitungannya seperti *accuracy*, *precision*, *recall*, *confusion matrix*, dan lain sebagainya

2.6. Deployment

Tahap *deployment* merupakan tahapan yang dilakukannya pembuatan laporan hasil analisis klasifikasi menggunakan model yang digunakan.

Pengumpulan data dilakukan dengan metode survei kepada beberapa keluarga. Data/atribut dalam penelitian ini diantaranya adalah Peran, Umur, Minat1, Minat2, Keahlian, Keahlian1, dan Aktivitas. Setelah data didapatkan dari survei oleh ± 80 keluarga, dihasilkan 287 *data record* yang dilakukan selama 3 hari. Hasil *data record* survei digambarkan pada Gambar 2.

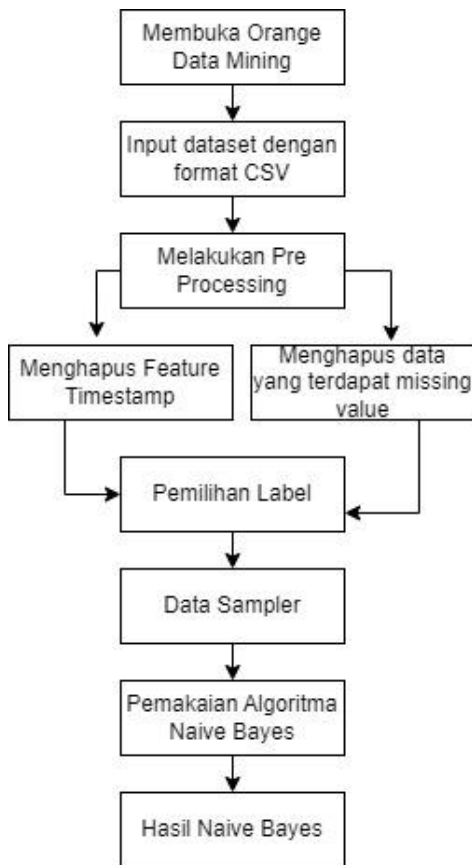
	Timestamp	Peran	Umur	Minat	Minat1	Minat2	Keahlian	Keahlian1	Aktivitas
0	21/05/2022 12:39:46	Anak	21-55	Alam	Hiburan	Animasi	Masak	Manajemen	Membersihkan Rumah
1	21/05/2022 12:44:19	Ayah	21-55	Makanan	Olahraga	Binatang	Swakarya/DIY	Berkebun	Membersihkan Rumah
2	21/05/2022 12:46:16	Anak	15<	Makanan	Animasi	Binatang	Swakarya/DIY	Sosialisasi	Belanja Kebutuhan
3	21/05/2022 12:49:24	Ibu	21-55	Makanan	Olahraga	Hiburan	Masak	Sosialisasi	Memasak
4	21/05/2022 13:03:39	Anak	15-20	Makanan	Hiburan	Keuangan	Masak	NaN	Membersihkan Rumah
...
282	25/05/2022 7:31:20	Anak	15<	Olahraga	Alam	Binatang	Berkebun	NaN	Mencuci Piring
283	25/05/2022 19:23:51	Anak	21-55	Makanan	Hiburan	Edukasi	Sosialisasi	NaN	Membersihkan Rumah
284	25/05/2022 19:28:45	Ibu	21-55	Makanan	Hiburan	Keuangan	Masak	Berkebun	Memasak
285	25/05/2022 19:29:51	Ayah	21-55	Teknologi	Hiburan	Edukasi	Otomotif	Kelistrikan	Memperbaiki Rumah/Kelistrikan/Perabotan
286	26/05/2022 17:43:45	Anak	15<	Makanan	Teknologi	Hiburan	Masak	Manajemen	Memasak

287 rows × 9 columns

Gambar 2. *Dataset* yang Dihasilkan dari Survei

Basis dari tahapan penelitian ini adalah dilakukannya metode CRISP-DM. Namun, sebelum digunakannya algoritma *naïve bayes*, perlu dilakukannya *preprocessing* seperti *Cleaning*. Proses *cleaning* pada penelitian ini digunakan untuk menghapus *feature timestamp* pada

dataset dan menghapus data yang terdapat *missing value*. Alur lengkap dari tahapan penelitian yang digunakan dijelaskan pada Gambar 3.



Gambar 3. Tahapan Pemodelan

3. HASIL DAN PEMBAHASAN

Terdapat beberapa aplikasi *text mining* yang dapat digunakan untuk perhitungan secara otomatis. Namun, penelitian ini memfokuskan perhitungan *text mining* dengan metode *Naive Bayes classifier* menggunakan aplikasi atau *tools Orange Data Mining*. Berikut hasilnya:

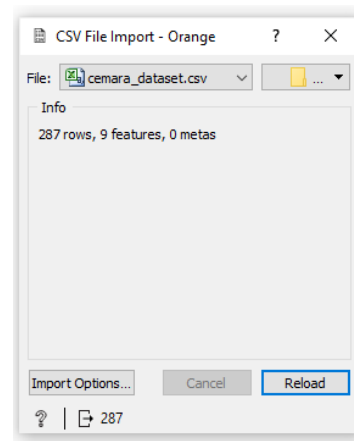
3.1. Persiapan *Dataset*

Hasil survei yang sudah dilakukan melalui *google form* menghasilkan 287 data yang akan dilakukan analisa. Di bawah ini terdapat beberapa tampilan dalam mempersiapkan *dataset*. Gambar 4 adalah *widget* untuk melakukan *import dataset* dengan format *csv*, Gambar 5 tampilan ketika *import file* dan Gambar 6 merupakan tampilan apabila *dataset* sudah di *import*



CSV File Import

Gambar 4. *Widget Import File*



Gambar 5. *Import File CSV*

Column type	1	2	3	4	5	6	7	8	9
1	Timestamp	Peran	Umur	Minat	Minat1	Minat2	Keshian	Keshian1	Aktivitas
2	21/05/2022 ...	Anak	21-35	Alam	Hiburan	Animasi	Masak	Manajemen	Membersihkan ...
3	21/05/2022 ...	Ayah	21-35	Makanan	Olahraga	Binatang	Swakarya/DIY	Berkebun	Membersihkan ...
4	21/05/2022 ...	Anak	15+	Makanan	Animasi	Binatang	Swakarya/DIY	Sosialisasi	Belanja ...
5	21/05/2022 ...	Ibu	21-35	Makanan	Olahraga	Hiburan	Masak	Sosialisasi	Memasak
6	21/05/2022 ...	Anak	15-20	Makanan	Hiburan	Keuangan	Masak	Kebijakan	Membersihkan ...
7	21/05/2022 ...	Ibu	21-35	Makanan	Alam	Kesehatan	Masak	Berkebun	Mengunjungi Anak ...
8	21/05/2022 ...	Ayah	21-35	Olahraga	Hiburan	Keuangan	Swakarya/DIY	Kebijakan	Mengunjungi ...
9	21/05/2022 ...	Anak	15+	Alam	Hiburan	Animasi	Swakarya/DIY	Sosialisasi	Mencuci Piring
10	21/05/2022 ...	Ibu	21-35	Olahraga	Hiburan	Keuangan	Masak	Berkebun	Memasak
11	21/05/2022 ...	Ayah	21-35	Teknologi	Hiburan	Animasi	Sosialisasi	Otomotif	Memperbaiki ...

Gambar 6. Tampilan *Dataset* di Orange

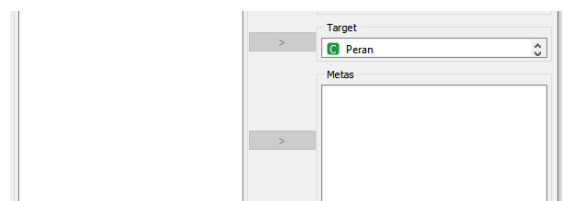
3.2. Pemilihan Label

Langkah yang dilakukan setelah *import file csv* adalah pemilihan label dan Gambar 7 merupakan tampilan dari *widget* pemilihan kolom yang akan dijadikan label. Dalam penelitian ini, peran merupakan *feature* yang dipilih untuk dijadikan label, dipaparkan melalui Gambar 8.



CSV File Import Select Columns

Gambar 7. *Widget Select Columns*



Gambar 8. Pemilihan Label

3.3. *Preprocessing*

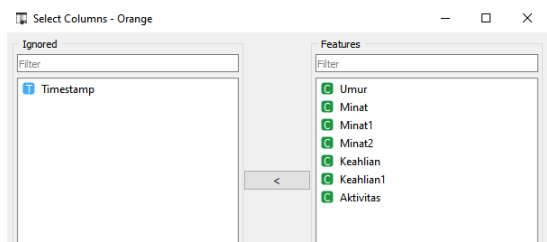
Setelah dilakukannya pemilihan *feature* untuk dijadikan label, tahap selanjutnya ialah proses *preprocessing* yaitu *cleaning*. Gambar 9 merupakan tampilan dari *widget preprocessing*.



Gambar 9. Widget Preprocessing

a) Hapus *feature timestamp*

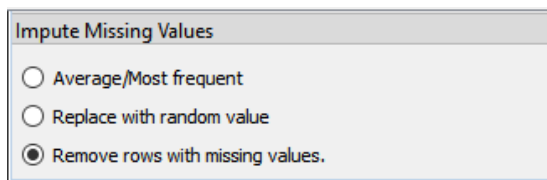
Dari total 9 *features* pada *dataset*, *feature* per sudah dipilih untuk menjadi label dan *feature timestamp* dipilih untuk dihapus dalam penelitian ini. Total *features* yang akan digunakan sekarang adalah 7 dan Gambar 10 menjelaskan mengenai apa saja yang menjadi *features* dan yang akan dihapus



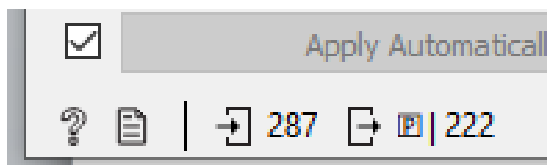
Gambar 10. Tampilan Hapus *Timestamp*

b) Hapus data yang terdapat *missing value*

Preprocessing dilanjut dengan penghapusan data/rows yang terdapat *missing value*, dijelaskan pada Gambar 11. Hasil dari penghapusan *missing value* digambarkan pada Gambar 12 yaitu terdapat 222 data yang akan di *split* untuk *training* dan *testing*



Gambar 11. Hapus *Missing Value*



Gambar 12. Hasil Data yang Sudah Dihapus

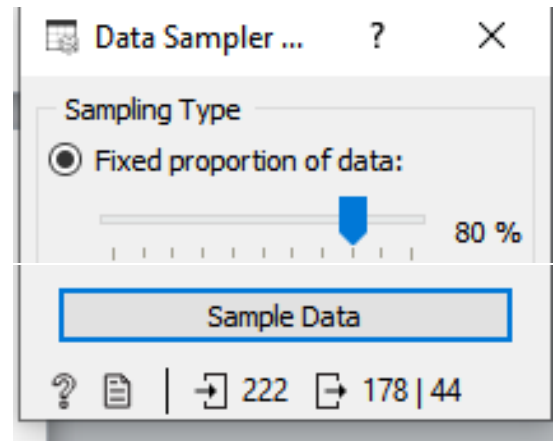
3.4. *Data Sampler*

Data sampler atau pemilihan data ini ialah proses pemilihan jumlah data yang akan dilatih (*data training*) dan jumlah data yang akan diuji (*data testing*). Seperti yang ditulis sebelumnya, proporsi dari *data training* dan *data testing* yaitu 80:20 dengan hasil 178 data yang

akan di *training* dan 44 data yang akan di *testing*. Gambar 13 merupakan tampilan *widget data sampler* dan Gambar 14 ialah tampilan untuk pengaturan proporsi dan hasil dari proporsi *data training* dan *data testing*.



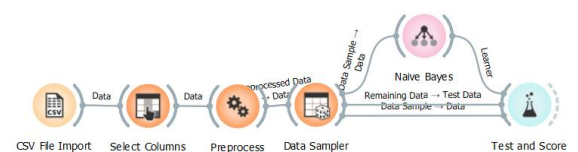
Gambar 13. *Widget Data Sampler*



Gambar 14. Proporsi *Data Train* dan *Test*

3.5. *Naïve Bayes*

Proporsi data sudah ditentukan dengan dilakukannya *split data*, maka langkah selanjutnya ialah melakukan pemodelan dengan *Naïve Bayes*. Gambar 15 merupakan tampilan secara utuh bagaimana pemodelan *Naïve Bayes* dilakukan pada *tools orange data mining*



Gambar 15. Pemodelan *Naïve Bayes*

3.6. Hasil

Evaluation results for target (None, show average over classes)					
Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.938	0.886	0.885	0.900	0.886

Evaluation results for target Anak					
Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.937	0.932	0.870	1.000	0.769

Gambar 16. Hasil Penelitian (a)

Evaluation results for target Ayah

Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.948	0.932	0.870	0.909	0.833

Evaluation results for target Ibu

Model	AUC	CA	F1	Precision	Recall
Naive Bayes	0.932	0.909	0.905	0.826	1.000

Gambar 17. Hasil Penelitian (b)

Hasil akurasi yang didapatkan dari penggunaan *tools orange data mining* menunjukkan 93% yang berarti mempunyai hasil yang baik. Selain akurasi, *precision* yang didapatkan pun mendapatkan hasil yang baik seperti yang ditampilkan pada Gambar 16 (a) dan (b).

Predictions - Orange

Naive Bayes	Peran	Umur	Minat	Minat1	Minat2	Keahlian	Keahlian1	Aktivitas
0.54 : 0.43 : 0.03 – Anak	Anak	21-55	Tekno...	Edukasi	Keuan...	Sosialis...	Manaje...	Members...
1.00 : 0.00 : 0.00 – Anak	Anak	15-20	Makanan	Tekn...	Hiburan	Masak	Sosialis...	Mencuci...
0.00 : 0.00 : 1.00 – Ibu	Ibu	21-55	Makanan	Olahr...	Keuan...	Masak	Berkebu...	Memasak
0.16 : 0.08 : 0.76 – Ibu	Anak	21-55	Hiburan	Keua...	Keseha...	Masak	Manaje...	Belanja K...
0.01 : 0.01 : 0.98 – Ibu	Ibu	21-55	Olahraga	Hibur...	Keseha...	Masak	Berkebu...	Memasak
0.03 : 0.00 : 0.97 – Ibu	Ibu	21-55	Makanan	Edukasi	Keuan...	Masak	Swakar...	Memasak
0.17 : 0.76 : 0.07 – Ayah	Ibu	21-55	Makanan	Hibur...	Keuan...	Sosialis...	Otomotif	Members...
1.00 : 0.00 : 0.00 – Anak	Anak	15-20	Tekno...	Hibur...	Animasi	Masak	Sosialis...	Mencuci...
0.04 : 0.84 : 0.12 – Ayah	Anak	21-55	Edukasi	Keua...	Keseha...	Sosialis...	Manaje...	Members...
0.90 : 0.09 : 0.01 – Anak	Anak	15-20	Makanan	Olahr...	Keseha...	Sosialis...	Otomotif	Members...

Gambar 18. Hasil Prediksi

Gambar 17 merupakan beberapa hasil dari prediksi *data testing* yang digunakan dalam penelitian. Pada tabel sebelah kiri yang bertuliskan *Naive Bayes*, merupakan tabel hasil prediksi dari *Naive Bayes* itu sendiri dan angka di data tersebut merupakan tingkat optimisme. Lalu pada sebelah kanan dari tabel *Naive Bayes* merupakan *data actual*.

		Predicted			
		Anak	Ayah	Ibu	Σ
Actual	Anak	10	1	2	13
	Ayah	0	10	2	12
	Ibu	0	0	19	19
Σ		10	11	23	44

Gambar 19. Confussion Matrix

Gambar 18 menampilkan hasil *confussion matrix* dari 44 *data testing* yang digunakan. Klasifikasi yang bernilai valid sejumlah 10 dari 10 dari data anak, 10 dari 11 dari data ayah, dan 19 dari 23 dari data ibu[7].

4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang sudah dievaluasi, penelitian ini menghasilkan model yang terbilang baik dengan hasil rata-rata akurasi (AUC) yang didapatkan sejumlah 93%, *class precision* “anak” didapatkan 100%,

class precision “ayah” didapatkan 90% dan *class precision* “ibu” didapatkan 82%.

Ucapan Terima Kasih

Puji syukur kepada Tuhan YME telah memberikan kelancaran sehingga penulis dapat menyelesaikan penelitian ini dan penulis mengucapkan terima kasih atas keterlibatan dosen maupun teman-teman dalam penelitian ini.

DAFTAR PUSTAKA

- [1] M. F. Andriansyah, D. Yusup, and A. Voutama, “Menggunakan Metode *Naive Bayes* Berbasis Website Web-Based Expert System of Covid-19 Early Detection Using *Naive Bayes* Method,” *J. Inf. Technol. Comput. Sci.*, vol. 4, no. 2, pp. 446–455, 2021.
- [2] M. T. Informatika and U. Amikom, “Analisis Pembobotan Kata pada Klasifikasi Text Mining,” vol. 3, no. 2, pp. 179–184, 2019.
- [3] A. V. Sudiantoro *et al.*, “Analisis Sentimen Twitter Menggunakan Text Mining Dengan Algoritma *Naive Bayes* Classifier,” *Din. Inform.*, vol. 10, no. 2, pp. 398–401, 2018.
- [4] D. M. Putri, M. T. Sari, and S. U. Khumaidatul, “Pengaruh Konflik Peran Ganda Terhadap Fear of Success Pada Wanita Karir yang Telah Menikah di Kota Samarinda,” *J. Psikol.*, vol. 9, no. 1, pp. 1–10, 2021.
- [5] Badan Pusat Statistik dan Kementerian Pemberdayaan Perempuan dan Perlindungan Anak, “Pr o f i l e r e m p u ON,” *Profil Peremp. Indones.*, 2019, [Online]. Available: https://www.kemenpppa.go.id/lib/uploads/list/b4bd-c-profil-perempuan-indonesial-_2019.pdf
- [6] L. Cahyadi and D. Prastyani, “Work Life Balance, Stres Kerja Dan Konflik Peran Terhadap Kepuasan Kerja Pada Wanita Pekerja,” *J. Ekon. J. Econ.*, vol. 11, no. 2, 2020, doi: 10.47007/jeko.v11i2.3576.
- [7] D. Dianti and W. Cahyati, “Persepsi Masyarakat Pada Program Studi Ilmu Komunikasi,” *Buana Komun. (Jurnal Penelit. dan Stud. Ilmu Komunikasi)*, vol. 2, no. 2, p. 116, 2022, doi: 10.32897/buanakomunikasi.2021.2.2.1134.
- [8] R. Safira, “Captain Marvel: Dominasi Maskulin Dalam Kesetaraan Gender Perempuan,” *Interak. Online*, vol. 10, no. 3, pp. 619–631, 2022, [Online]. Available: <https://fisip.undip.ac.id/>
- [9] M. Prasojo and J. Triwidianti, “Prediksi Prestasi Siswa SMK Masuk Pasar Kerja Menggunakan Teknik Data Mining (Studi Kasus SMKN 1 Kota Agung Timur Tanggamus, Lampung),” *Semin. Nas. Has. Penelit. dan Pengabd. Masy.*, pp. 134–150, 2021.
- [10] M. M. Yin, M. R. Mahmoudi, and A. Abbasalizadeh, “Analysis of mystical concepts in

- Khaghani's Divan," *Digit. Scholarsh. Humanit.*, vol. 35, no. 2, pp. 485–491, 2020, doi: 10.1093/LLC/FQZ015.
- [11] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.
- [12] W. P. Nurmayanti, "Penerapan Naive Bayes dalam Mengklasifikasikan Masyarakat Miskin di Desa Lepak," *Geodika J. Kaji. Ilmu dan Pendidik. Geogr.*, vol. 5, no. 1, pp. 123–132, 2021, doi: 10.29408/geodika.v5i1.3430.
- [13] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode *Naive Bayes* Classifier Berbasis Konversi Ikon Emosi," *J. SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.
- [14] M. Nizam Fadli, I. Sudahri Damanik, E. Irawan, S. Tunas Bangsa, and S. Utara, "KLIK: Kajian Ilmiah Informatika dan Komputer Penerapan Metode Naive Bayes Dalam Menentukan Tingkat Kenyamanan Pada Rumah Sakit Terhadap Pasien," *Media Online*, vol. 2, no. 3, pp. 117–122, 2021, [Online]. Available: <https://djournals.com/klik>
- [15] D. Wilandani and Purwanto, "Penerapan Algoritma Naive Bayes dalam Mengklasifikasikan Media Sosial untuk Mengamati Trend Kuliner," vol. 8, no. 1, pp. 31–39, 2022.
- [16] S. I. Nurhafida and F. Sembiring, "Analisis Text Clustering Masyarakat di Twiter Mengenai Mcdonald'sxbs Menggunakan Orange Data Mining," *Semin. Nas. Sist. Inf. dan Manaj. Inform.*, vol. 1, no. 1, pp. 28–35, 2021.
- [17] A. Voutama, I. Maulana, and N. Ade, "Interactive M-Learning Design Innovation using Android-Based Adobe Flash at WFH (Work From Home)," *Sci. J. Informatics*, vol. 8, no. 1, pp. 127–136, 2021, doi: 10.15294/sji.v8i1.27880.



ANALISIS SENTIMEN ULASAN APLIKASI MYPERTAMINA PADA *GOOGLE PLAY STORE* MENGGUNAKAN ALGORITMA NBC

Rihan Maulana¹, Apriade Voutama², Taufik Ridwan³

^{1,2,3} Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang
Karawang, Jawa Barat, Indonesia 41361

rihan.maulana19028@student.unsika.ac.id, apriade.voutama@staff.unsika.ac.id, taufik.ridwan@cs.unsika.ac.id

Abstract

The MyPertamina application is a digital service application from Pertamina with various services initiated and created to carry out purchases in vehicle fuel transactions. The existence of the MyPertamina application made by PT Pertamina caused reactions and criticism from some application users. Each user has various opinions about the MyPertamina application, as evidenced by the different star ratings in Google Play Store reviews. This study aims to determine the sentiments of MyPertamina application users, which are divided into two classes, namely positive and negative. The dataset in this study uses scraping results from user reviews on the Google Play Store. The data taken to carry out sentiment analysis is from 1 July 2022 to 31 July 2022, and the dataset is taken randomly. The dataset is classified by rating: ratings 4 and 5 as positive sentiment and ratings 1, 2, and 3 as negative sentiment. In this study, Google Colab tools will be used using Python programming. The dataset used is 5722 data labeled positive and negative, with the division of training data by 80% and test data by 20%. The MyPertamina application sentiment analysis results tend to be hostile towards using the application. This study used classification with the Naïve Bayes Classifier (NBC) algorithm. By using the Naïve Bayes Classifier algorithm, it produces 87% accuracy, 86% precision, 90% recall, and 87% f1-score.

Keywords: Classification Method, MyPertamina, Naïve Bayes Classifier, Sentiment Analysis, Text Mining

Abstrak

Aplikasi MyPertamina merupakan aplikasi layanan digital dari Pertamina dengan berbagai layanan yang digagas dan diciptakan untuk melakukan pembelian dalam bertransaksi bahan bakar kendaraan. Adanya aplikasi MyPertamina yang dibuat oleh PT Pertamina menimbulkan reaksi dan kritik dari beberapa pengguna aplikasi. Setiap pengguna memiliki berbagai opini terhadap aplikasi MyPertamina dibuktikan dengan peringkat bintang yang beragam dalam ulasan *Google Play Store*. Tujuan dari penelitian ini adalah untuk mengetahui sentimen pengguna aplikasi MyPertamina yang dibagi ke dalam dua kelas yaitu positif dan negative. *Dataset* dalam penelitian ini menggunakan hasil *scraping* dari ulasan pengguna pada *Google Play Store*. Data yang diambil untuk melakukan analisis sentimen yaitu pada rentang tanggal 1 Juli 2022 - 31 Juli 2022 dan *dataset* diambil dengan acak. *Dataset* tersebut dengan pengklasifikasian *rating* yaitu *rating* 4 dan 5 sebagai sentimen positif, *rating* 1, 2 dan 3 sebagai sentimen negatif. Pada penelitian kali ini akan menggunakan *tools Google Colab* dengan menggunakan bahasa pemrograman *Python*. *Dataset* yang digunakan sebanyak 5722 data berlabel positif dan negatif dengan pembagian data latihan sebanyak 80% dan data uji sebanyak 20%. Hasil analisis sentimen aplikasi MyPertamina cenderung bersifat *negative* terhadap penggunaan aplikasi. Penelitian ini menggunakan klasifikasi dengan algoritma *Naïve Bayes Classifier* (NBC). Dengan menggunakan algoritma *Naïve Bayes Classifier* menghasilkan *accuracy* 87 %, *precision* 86%, *recall* 90%, dan *f1-score* 87%.

Kata kunci: Analisis Sentimen, Metode Klasifikasi, MyPertamina, *Naïve Bayes Classifier*, *Text Mining*

1. PENDAHULUAN

Aplikasi MyPertamina merupakan aplikasi layanan digital dari Pertamina dengan berbagai layanan yang ada di dalamnya yaitu untuk menemukan lokasi SPBU Pertamina terdekat, pembayaran digital yang sekaligus mendapatkan *loyalty point* dan sistem pencatatan ketika belanja bensin

bulanan agar lebih mudah. Aplikasi ini digagas dan diciptakan oleh PT Pertamina sebagai solusi bagi para pengguna BBM supaya memberikan kemudahan, cepat serta aman dalam melaksanakan pembelian dalam bertransaksi bahan bakar kendaraan [1]. Dampak dari naiknya minyak dunia membuat pemerintah Indonesia

berencana untuk membatasi penggunaan bahan bakar bersubsidi yaitu pertalite dan solar. Pada tanggal 1 Juli 2022, pemerintah melalui PT Pertamina Persero mewajibkan pengguna BBM bersubsidi untuk mendaftar melalui aplikasi MyPertamina atau *website* Pertamina. Kebijakan yang dilakukan oleh pemerintah bertujuan agar subsidi BBM tepat sasaran, karena data yang didaftarkan di aplikasi akan menunjukkan konsumen mana yang berhak mendapatkan subsidi atau tidak.

Keberadaan aplikasi MyPertamina yang dibuat oleh PT Pertamina Persero menimbulkan reaksi dan kritikan dari beberapa pengguna aplikasi. Setiap pengguna memiliki berbagai opini terhadap aplikasi MyPertamina yang baru saja di sosialisasikan oleh pemerintah. Kebijakan yang dilakukan pemerintah menimbulkan pro dan kontra di masyarakat, konsumen akhirnya menuliskan *rating* dan *review*-nya pada *Google Play Store* dimana aplikasi tersebut dapat diunduh.

Pada masalah yang muncul dari melihat persoalan yang ditemukan, maka diperlukan analisis sentimen terhadap *review* atau ulasan dari warga yang menggunakan aplikasi MyPertamina sehingga informasi sentimen dapat diketahui oleh pemerintah terkait kebijakan yang telah disosialisasikan kepada pengguna BBM bersubsidi. Analisis Sentimen diperlukan agar pemerintah dapat menganalisis pendapat, melakukan evaluasi kebijakan, dan penilaian dari para pengguna aplikasi. Dalam hal ini, *review* yang terdapat di aplikasi tersebut kurang baik, terbukti dengan *rating* yang diberikan oleh pengguna aplikasi. Jumlah data ulasan yang besar membuat sulit untuk membedakan antara sentimen yang bersifat positif dan negatif sehingga perlu dilakukan proses analisis terhadap *review* pada aplikasi tersebut. Data yang akan digunakan yaitu hasil *scraping* data *rating* dan ulasan di *Google Play Store* pada aplikasi MyPertamina.

Pada penelitian sebelumnya yang dilaksanakan oleh [2] penelitian tersebut menggunakan algoritma *Naïve Bayes Classifier* memperoleh hasil penelitian dengan nilai akurasi sebesar 85% dengan nilai akurasi yang baik maka penelitian ini menggunakan algoritma *Naïve Bayes* dengan membedakan ulasan dari aplikasi yang berbeda. Selain itu, pada penelitian yang ditulis oleh [3] menyarankan agar menggunakan *dataset* yang besar dan tidak terlalu jauh variasi perbandingan komentar antara positif dan negatif. Maka, penelitian kali ini akan menggunakan *dataset* yang lebih banyak dan variasi komentar yang berimbang.

Dari ulasan di atas, perlu dilakukan analisis dengan metode tertentu untuk meringkas ulasan. Setiap ulasan yang dikumpulkan akan mendapatkan kesimpulan dan dilakukannya analisis sentimen agar pengalaman pengguna dapat diperoleh dari aplikasi MyPertamina untuk periode tertentu. Untuk itu pada penelitian ini akan menggunakan algoritma NBC. *Naïve Bayes Classifier* merupakan metode yang mengandalkan probabilitas untuk mengklasifikasikan teks. Algoritmanya sederhana namun memiliki akurasi dan

presisi yang tinggi merupakan salah satu kelebihan dari algoritma NBC. Metode ini akan digunakan pada analisis sentimen aplikasi MyPertamina berdasarkan ulasan dari *google play store* untuk menentukan kualitas dari layanan aplikasi berdasarkan hasil ulasan pengguna. Algoritma *Naïve Bayes Classifier* digunakan karena algoritma ini menemukan nilai probabilitas tertinggi untuk mengklasifikasikan data uji ke dalam kategori yang paling sesuai sehingga merupakan algoritma yang banyak digunakan dan sesuai untuk penelitian kali ini.

Tujuan dari penelitian ini adalah untuk mengetahui sentimen pengguna aplikasi MyPertamina yang dibagi ke dalam dua kelas yaitu positif dan negatif sehingga hasil opini pengguna dapat menjadi evaluasi bagi pemerintah khususnya Pertamina dalam melakukan kebijakan dan *update* penggunaan aplikasi. Setelah hasil sudah didapatkan, kemudian hasil klasifikasi akan dievaluasi *accuracy*, *precision*, *recall* dan *f1-score* dari pemodelan yang menggunakan algoritma *Naïve Bayes*. Diharapkan dari hasil penelitian ini dapat menjadi kontribusi kepada pemerintah dan PT Pertamina Persero dalam mengambil langkah dan kebijakan yang tepat dalam penggunaan atau pengembangan aplikasi.

Aplikasi adalah suatu program pada sistem komputer yang digunakan orang untuk melakukan sesuatu. Aplikasi merupakan suatu komponen yang bermanfaat sebagai media yang digunakan untuk melakukan pengelolaan data-data maupun kegiatan lainnya [4]. Sistem aplikasi *mobile* merupakan aplikasi yang penggunanya dapat lancar berpindah-pindah dari satu tempat ke suatu tempat lain dengan tidak terjadi komunikasi yang terputus atau terhenti [5]. Aplikasi *Mobile* adalah aplikasi yang dirancang untuk *platform* khusus berbasis *mobile* yang bertujuan membantu *user* sesuai dalam jenis aplikasi yang disediakan [6].

Text Mining adalah suatu proses penambangan inti sari dari suatu dokumen data berupa teks yang bentuknya lebih tidak teratur yang dapat dibuat pola untuk menghasilkan sebuah informasi yang berguna. Proses penambangan data dapat bersumber melalui *website*, yang datanya dapat berupa data dokumen berbentuk teks dari *website* yang di dalamnya berisi ulasan, opini, kritik dan *feedback*. *Text mining* merupakan salah satu ragam dari *data mining* yang bertujuan untuk mengambil format yang dapat diolah dari sekumpulan data-data yang berupa *text* dalam jumlah besar [7]. Sebelum *text mining* dilakukan terdapat tahapan sebelumnya yaitu *text pre-processing* yaitu *case folding*, *cleansing*, *tokenizing*, *stemming*, dan *filtering* yang menggunakan *stopword removal* [8].

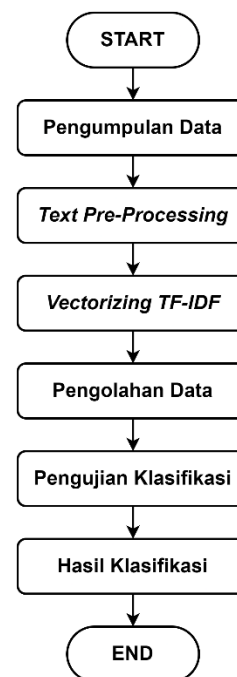
Analisis Sentimen merupakan teknik penggalian informasi yang menganalisis penilaian, pendapat, opini, sikap, aksi dan emosi terhadap suatu entitas seperti isu, layanan, produk, atau masalah tertentu [9]. Analisis Sentimen penting dilakukan dengan perkembangan yang pesat suatu opini di media sosial dan berbagai diskusi di ulasan atau

suatu *review* produk dan layanan [10]. Karena pertumbuhan era digital yang sangat pesat membuat masyarakat pada zaman modern tidak dapat lepas dari penggunaan teknologi, salah satunya yaitu media sosial membuat orang mengungkapkan dan menuliskan pendapat yang tertulis dapat melalui media sosial atau melalui suatu ulasan sehingga analisis sentimen dapat dilakukan [11].

Algoritma *Naïve Bayes* adalah algoritma yang mengandalkan kemungkinan atau probabilitas dengan pengklasifikasi statistik kelas data, klasifikasi ini dikelompokkan ke kelas tertentu sesuai teori probabilitas [12]. Algoritma *Naïve Bayes* dapat menarik kesimpulan berdasarkan klasifikasi *data training* yang disimpan sebelumnya. Meskipun sifat independen antara kata (istilah) atau parameter dalam dokumen tidak sepenuhnya terpenuhi dengan sempurna, tapi kinerja *Naïve Bayes* untuk klasifikasi relatif dapat diandalkan dan bahkan lebih unggul dalam hal realistis, kecepatan tinggi dan akurasi yang baik [13]. Metode klasifikasi dapat menemukan pola suatu model dengan membedakan suatu konsep atau kelas yang tidak diketahui namanya dan termasuk ke dalam metode *data mining*. Metode klasifikasi bersifat *supervised learning* dan juga didefinisikan sebagai metode *data mining* yang membagi data ke dalam kelas-kelas tertentu berdasarkan kategorinya. Proses klasifikasi terdiri dari empat komponen dasar yaitu *class*, *predictor*, *training set* dan *data training* [14]. *Naïve Bayes* ini memprediksi probabilitas peluang di masa mendatang bersumber pada pengalaman yang sudah ada sebelumnya [15]. Algoritma ini dapat memungkinkan untuk melakukan pekerjaan yang lebih bagus dari kejadian nyata yang rumit dan data yang kompleks jauh dari yang diharapkan.

2. METODE PENELITIAN

Penelitian yang dilakukan akan dilakukan dengan berbagai tahapan, secara umum dengan melakukan proses pengumpulan data (*scraping*), *text pre-processing* (*case folding*, *cleansing*, *tokenizing*, *stemming*, dan *filtering*), pembobotan kata (*vectorizing TF-IDF*), pengolahan data, pengujian klasifikasi menggunakan algoritma NBC lalu kemudian hasil klasifikasi [16]. Pada penelitian kali ini akan menggunakan *tools Google Colab* dengan menggunakan bahasa pemrograman *Python*. Dukungan *library* menjadi salah satu penggunaan *Python* dalam proses analisis *sentiment* semakin mudah. *Library* yang digunakan dalam proses perhitungan harus di-instal terlebih dahulu. *Library* ini termasuk *pandas*, *numpy*, *matplotlib*, *sklearn*, dan *sastrawi*. Diagram alur tahapan penelitian dapat dilihat pada Gambar 1 berikut.



Gambar 1. Diagram Alur Metode Penelitian

2.1 Pengumpulan Data

Pengambilan data pada penelitian kali ini menggunakan hasil *scraping* data. Data yang ingin diolah adalah *rating* dan ulasan dari pengguna aplikasi MyPertamina di *Google Play Store*. *Rating* tersebut akan digunakan untuk memilah sentimen positif dan negatif. Data yang diambil untuk melakukan analisis sentimen yaitu pada rentang tanggal 1 Juli 2022 sampai tanggal 31 Juli 2022 dan *dataset* diambil dengan acak. Pengklasifikasian *rating* sebagai berikut yaitu *rating* 4 dan 5 sebagai sentimen positif, *rating* 1, 2 dan 3 sebagai sentimen negatif. Hasil *scraping* yang diambil pada ulasan setelah melakukan proses pengolahan hasil data yang telah dihimpun dan dilakukan *cleansing* data dapat dilihat pada Tabel 1 berikut.

Tabel 1. Contoh *Dataset* Ulasan Review

<i>Comment</i>	<i>Value</i>
Sudah daftar keluar terus abis abisin kouta aja	NEGATIVE
Sangat membantu kelancaran serta kemudahan dalam bertransaksi	POSITIVE
Aplikasinya masih banyak eror	NEGATIVE
Langkah upaya subsidi tepat sasaran	POSITIVE
Mempersulit rakyat aja buat apa coba pakai beginian terus yang ga punya android gimana??? Motor gonta ganti gimana, kacau!!!!!!	NEGATIVE

2.2 Text Pre-processing

Text Pre-processing merupakan tahap pertama dalam pemrosesan data supaya data dapat diolah dan data siap untuk proses penelitian salah satunya analisa sentimen. Proses *text pre-processing* dapat berupa *cleaning*, *case-folding*, *tokenizing*, *stemming*, dan *filtering* [3].

a. Casefolding

Casefolding merupakan tahapan yang mengubah suatu kata atau huruf pada ulasan menjadi huruf yang kecil semua atau menjadi huruf besar semua. Pada implementasi *casefolding* untuk penelitian yang sekarang mengubah semua tulisan menjadi huruf kecil semua. Contoh proses *casefolding* dapat dilihat pada Tabel 2 berikut.

Tabel 2. Tahap *Casefolding* Ulasan

Sebelum <i>Casefolding</i>	Sesudah <i>Casefolding</i>
udah daftar tinggal tunggu di verifikasi nomor Hp, kode otp udah dikirim otomatis masa gabisa ke verifikasi sih? Gimana aneh banget gada tombol apa" Malah balik ke menu utama sedangkan akun sudah jadi tapi belum terverifikasi juga kan aneh malah gabisa login	udah daftar tinggal tunggu di verifikasi nomor hp, kode otp udah dikirim otomatis masa gabisa ke verifikasi sih? gimana aneh banget gada tombol apa" malah balik ke menu utama sedangkan akun sudah jadi tapi belum terverifikasi juga kan aneh malah gabisa login

b. Tokenizing

Tokenizing merupakan tahapan yang yang membuat setiap kata menjadi berurutan dari suatu kalimat, setiap kata akan di pisah-pisah atau dilakukan separasi. Contoh proses *tokenizing* dapat dilihat pada Tabel 3 berikut.

Tabel 3. Tahap *Tokenizing* Ulasan

Sebelum <i>Tokenizing</i>	Sesudah <i>Tokenizing</i>
udah daftar tinggal tunggu di verifikasi nomor hp, kode otp udah dikirim otomatis masa gabisa ke verifikasi sih? gimana aneh banget gada tombol apa" malah balik ke menu utama sedangkan akun sudah jadi tapi belum terverifikasi juga kan aneh malah gabisa login	['udah', 'daftar', 'tinggal', 'tunggu', 'di', 'verifikasi', 'nomor', 'hp', 'kode', 'otp', 'udah', 'dikirim', 'otomatis', 'masa', 'gabisa', 'ke', 'verifikasi', 'sih', 'gimana', 'aneh', 'banget', 'gada', 'tombol', 'apa', 'malah', 'balik', 'ke', 'menu', 'utama', 'sedangkan', 'akun', 'sudah', 'jadi', 'tapi', 'belum', 'terverifikasi', 'juga', 'kan', 'aneh', 'malah', 'gabisa', 'login']

c. Filtering

Filtering merupakan tahapan yang digunakan untuk menghapus kata yang kurang bermanfaat atau kurang bermakna untuk proses dalam melakukan analisis sentimen. Dalam penelitian ini, akan menggunakan metode *stopwords* yaitu kata-kata yang tidak penting akan dibuang berdasarkan data yang diproses ada dalam metode. Contoh dari *stopwords* yaitu kata "di", "dan", "kan", "yang" dan sebagainya. Contoh proses *filtering* dapat dilihat pada Tabel 4 berikut.

Tabel 4. Tahap *Filtering* Ulasan

Sebelum <i>Filtering</i>	Sesudah <i>Filtering</i>
['udah', 'daftar', 'tinggal', 'tunggu', 'di', 'verifikasi', 'nomor', 'hp', 'kode', 'otp', 'udah', 'dikirim', 'otomatis', 'masa', 'gabisa', 'ke', 'verifikasi', 'sih', 'gimana', 'aneh', 'banget', 'gada', 'tombol', 'apa', 'malah', 'balik', 'ke', 'menu', 'utama', 'sedangkan', 'akun', 'sudah', 'jadi', 'tapi', 'belum', 'terverifikasi', 'juga', 'kan', 'aneh', 'malah', 'gabisa', 'login']	['udah', 'daftar', 'tinggal', 'tunggu', 'verifikasi', 'nomor', 'hp', 'kode', 'otp', 'udah', 'dikirim', 'otomatis', 'masa', 'gabisa', 'verifikasi', 'gimana', 'aneh', 'banget', 'gada', 'tombol', 'apa', 'malah', 'balik', 'menu', 'utama', 'sedangkan', 'akun', 'sudah', 'jadi', 'belum', 'terverifikasi', 'aneh', 'malah', 'gabisa', 'login']

d. Stemming

Stemming merupakan tahapan yang digunakan untuk mengkonversikan kata-kata yang berimbuhan menjadi kata dasar atau menghilangkan imbuhan. Menggunakan *library* sastrawi yang digunakan sebagai acuan ketentuan bahasa Indonesia, maka keluaran hasil *stemming* akan sesuai dengan ketentuan *library* sastrawi. Contoh proses *stemming* dapat dilihat pada Tabel 5 berikut.

Tabel 5. Tahap *Stemming* Ulasan

Sebelum <i>Stemming</i>	Sesudah <i>Stemming</i>
['udah', 'daftar', 'tinggal', 'tunggu', 'di', 'verifikasi', 'nomor', 'hp', 'kode', 'otp', 'udah', 'dikirim', 'otomatis', 'masa', 'gabisa', 'ke', 'verifikasi', 'sih', 'gimana', 'aneh', 'banget', 'gada', 'tombol', 'apa', 'malah', 'balik', 'ke', 'menu', 'utama', 'sedangkan', 'akun', 'sudah', 'jadi', 'tapi', 'belum', 'terverifikasi', 'juga', 'kan', 'aneh', 'malah', 'gabisa', 'login']	['udah', 'daftar', 'tinggal', 'tunggu', 'verifikasi', 'nomor', 'hp', 'kode', 'otp', 'udah', ' kirim', 'otomatis', 'masa', 'gabisa', 'verifikasi', 'gimana', 'aneh', 'banget', 'gada', 'tombol', 'malah', 'balik', 'menu', 'utama', 'sedang', 'akun', 'sudah', 'belum', 'verifikasi', 'aneh', 'malah', 'gabisa', 'login']

2.3 Vectorizing TF-IDF

Pembobotan kata dengan menggunakan *vectorizing TF-IDF* dilaksanakan dengan cara memberikan bobot pada setiap kata dari masing-masing dokumen/kalimat dengan keseringan kata yang muncul terbanyak akan ditampilkan terlebih dahulu yang sudah melalui tahap *pre-processing*. Ada dua proses yang digunakan dalam tahapan pembobotan kata yaitu tahap *Term Frequency* (TF) dan tahap *Inverse Document Frequency* (IDF). *Term Frequency* (TF) merupakan suatu term dalam dokumen terkait yang dihitung frekuensi kemunculannya. Bobot kata akan semakin tinggi apabila banyak jumlah kata yang sering muncul (term) dalam suatu dokumen semakin banyak. *Inverse Document Frequency* (IDF) merupakan hubungan ketersediaan sebuah term dalam seluruh dokumen yang menunjukkan sebuah perhitungan. Besar nilai IDF akan semakin besar apabila besaran dokumen yang didalamnya terdapat suatu term

tertentu semakin sedikit [2]. Persamaan berikut untuk menghitung TF-IDF bobot setiap kata.

$$tf - idf_{t,d} = tf_d * idf_t \tag{1}$$

2.4 Pengolahan Data

Pengolahan data pada penelitian ini akan menggunakan algoritma *Naïve Bayes Classifier*. Algoritma NBC adalah klasifikasi probabilitas sederhana dengan menggunakan model statistik untuk menghitung peluang kelas yang memiliki setiap kelompok dengan atribut yang tampil, dimana label kelas diambil dari beberapa himpunan berhingga. Metode *Naïve Bayes* merupakan metode yang terdiri atas beberapa keunggulan, seperti perhitungan sederhana, presisi yang tinggi, dan memiliki kecepatan apabila melakukan pemrosesan *database* yang banyak dan besar [8]. Rumus *Naïve Bayes* memiliki bentuk umum yang sebagai berikut :

$$P(c|X) = \frac{P(x|c)P(c)}{P(x)} \tag{2}$$

Keterangan :

- c : Hipotesis data X merupakan suatu class spesifik
- X : Data dengan class yang belum diketahui
- P(c|X) : Probabilitas hipotesis berdasarkan kondisi X (*posteriori probability*)
- P(x|c) : Probabilitas hipotesis x berdasarkan kondisi pada hipotesis c
- P(x) : Probabilitas c
- P(c) : Probabilitas hipotesis c (*prior probability*)

2.5 Pengujian Klasifikasi *Confusion Matrix*

Confusion matrix merupakan suatu matriks pengukuran performa yang berisi nilai perkiraan benar dan nilai perkiraan salah dimana keluarannya dapat berupa dua kelas atau lebih. Nilai *recall*, *precision*, *accuracy* dan *f1-score* akan menjadi evaluasi dalam pengujian keakuratan hasil pencarian nilai dalam penelitian. Struktur *Confusion matrix* akan menghasilkan angka yang terdapat pada Tabel 6 berikut.

Tabel 6. *Confusion Matrix*

	Actual Positive (1)	Actual Negative (0)
Predicted Positive (1)	TP	FP
Predicted Negative (0)	FN	TN

Tabel 6 menjelaskan 4 sisi nilai dari *confusion matrix* terdiri dari TP merupakan singkatan dari *True Positive*, yaitu angka perkiraan positif sesuai dengan nilai aktual benar. FP merupakan singkatan dari *False Positive*, yaitu angka perkiraan positif tidak sesuai dengan nilai aktual salah. FN merupakan singkatan dari *False Negative*, yaitu angka perkiraan negatif tidak sesuai dengan nilai aktual benar. Dan

TN merupakan singkatan dari *True Negative*, yaitu angka perkiraan negatif sesuai dengan nilai aktual salah.

Accuracy dapat memprediksi nilai dengan benar sehingga menggambarkan seberapa akurat dan tepat (*machine learning*). Dengan kata lain, merupakan perbandingan rasio kasus yang diidentifikasi benar (*negative* dan *positive*) dengan semua data yang ada. Rumus *accuracy* adalah :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

Precision mengevaluasi kemampuan tingkat ketelitian data antara data aktual dengan hasil prediksi dalam model untuk menemukan peringkat yang paling relevan. Rumus *precision* adalah :

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

Recall untuk memaparkan keberhasilan model dalam mendapatkan kembali sebuah informasi. Sehingga *recall* untuk mengevaluasi kemampuan sistem untuk menemukan informasi yang sesuai dari semua data yang benar positif. Rumus *recall* adalah :

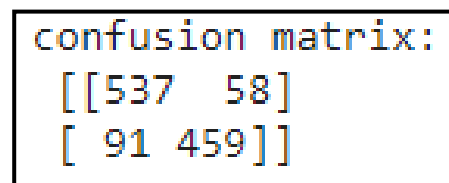
$$Recall = \frac{TP}{TP+FN} \tag{5}$$

F1-Score adalah komparasi perhitungan rata-rata nilai dari *recall* dan *precision* yang telah di bobotkan.

$$F1\ Score = 2 \times \frac{Recall \times Precision}{Recall+Precision} \tag{6}$$

3. HASIL DAN PEMBAHASAN

Data ulasan yang telah diambil datanya dengan melakukan *scraping* di aplikasi MyPertamina pada platform *Google Play Store*. *Dataset* yang akan digunakan dalam penelitian kali ini yaitu ulasan *random sampling* pada rentang waktu tanggal 1 Juli 2022 sampai 31 Juli 2022, terdiri dari 5722 data yang didalamnya terdapat data dari data ulasan negatif sebanyak 2861 dan data ulasan positif sebanyak 2861. *Dataset* ulasan lalu dipecah menjadi dua bagian yaitu 80% data latih sebanyak 4577 dan 20% data uji sebanyak 1145. *Dataset* terdiri dari 2 kolom yang berisi ulasan dari pengguna dan kolom yang menjadi value label yaitu *positive* dan *negative* yang diberikan pada setiap komentar.



Gambar 2. Hasil *Confusion Matrix*

Hasil proses *confusion matrix* dapat dilihat pada Gambar 2. Hasil penelitian yang diukur dengan *confusion matrix* dari data uji sebanyak 1145 ulasan, sistem telah berhasil melakukan klasifikasi sebanyak 628 ulasan komentar positif

dan 517 ulasan komentar negatif. Hasil data tersebut dengan rincian sebanyak 628 ulasan positif diprediksi benar sebanyak 537 ulasan dan sebanyak 91 ulasan diprediksi salah. Kemudian, sebanyak 517 ulasan negatif diprediksi benar sebanyak 459 ulasan dan sebanyak 58 ulasan diprediksi salah.

Penelitian kali ini akan menjelaskan hasil yang didapat dan menguraikan performansinya dengan data nilai *confusion matrix* yang telah didapatkan. Hasil *accuracy*, *precision*, *recall* dan *f1-score* dapat dihitung data dalam *confusion matrix*.

Tabel 7. Hasil *Naïve Bayes Classifier*

Performa <i>Confusion Matrix</i> binary class			
<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
87 %	86 %	90 %	87 %

Berdasarkan hasil uji yang didapatkan, pada Tabel 7 dengan menggunakan algoritma *Naïve Bayes Classifier* menghasilkan hasil *accuracy* sebesar 87 %, hasil *precision* sebesar 86%, hasil *recall* sebesar 90% dan hasil *f1-score* sebesar 87%.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilaksanakan pengujian pada sistem, maka dapat disimpulkan sebagai berikut :

1. Secara keseluruhan dengan menggunakan algoritma *Naïve Bayes* dalam mengklasifikasikan ulasan komentar perihal aplikasi MyPertamina pada *Google Play Store* dapat melakukan klasifikasi dengan sangat baik untuk memberikan label value *positive* dan *negative* dengan akurasi yang mencapai 87%.
2. Dataset dengan menggunakan data yang seimbang antara label positif dan label negatif, hasil klasifikasi yang dihasilkan cukup baik dan akurat dilihat dari akurasi yang mencapai 87% menggunakan algoritma *Naïve Bayes*.
3. Algoritma yang diterapkan pada *dataset* dari penelitian, algoritma NBC dapat memprediksi data *review* aplikasi MyPertamina dengan data ulasan negatif dengan nilai 90% *true positive*.

Dataset yang dijadikan objek penelitian, dilihat dari data yang berdasarkan label positif dan negatif ulasan aplikasi, hasil analisis sentimen aplikasi MyPertamina cenderung bersifat negatif. Analisis *sentiment* negatif yaitu penggunaan aplikasi yang masih mengeluhkan penggunaan aplikasi yang belum siap untuk diakses untuk melakukan pembelian BBM. Terdapat *error* ketika proses pendaftaran seperti kode otp yang tidak terkirim ke nomor yang didaftarkan. Akan tetapi, ada juga pengguna yang mendukung dan memberi ulasan positif untuk aplikasi MyPertamina yang berfungsi untuk membatasi para konsumen yang tidak layak mendapatkan subsidi.

Pemerintah dapat memperbaiki sistem aplikasi pendaftaran dari MyPertamina dikarenakan banyaknya respon masyarakat mengenai kesulitan saat proses pendaftaran dan pengiriman kode otp serta lebih siap ketika aplikasi digunakan secara bersamaan agar tidak terjadi susah masuk ke *system* sehingga program ini dapat terus berjalan dan penerima BBM bersubsidi tepat sasaran.

Penelitian selanjutnya, disarankan untuk melakukan perbandingan terhadap algoritma lain, tidak menggunakan algoritma *Naïve Bayes Classifier* untuk melakukan analisis tetapi dapat menggunakan algoritma lain seperti K-NN dan *Decision Tree* sehingga melakukan perbandingan yang dapat diketahui mana algoritma yang terbaik untuk melakukan analisis sentimen.

DAFTAR PUSTAKA

- [1] F. M. Oktaviana, D. Wijayanto, and T. Wahyudi, "Pengaruh Social Marketing Campaign Terhadap Keputusan Konsumen Bertransaksi Menggunakan App Mypertamina Di Pontianak," *J. TIN*, vol. 5, no. 1, pp. 23–29, 2021.
- [2] I. Saputra, T. Djatna, R. R. A. Siregar, D. A. Kristiyanti, H. R. Yani, and A. A. Riyadi, "Text Mining of PeduliLindungi Application Reviews on Google Play Store," *Fakt. Exacta*, vol. 15, no. 2, pp. 101–108, 2022.
- [3] G. K. Locarso, "Analisis Sentimen Review Aplikasi Pedulilindungi pada Google Play Store Menggunakan NBC," *J. Tek. Inform. Kaputama*, vol. 6, no. 2, 2022.
- [4] A. Voutama, "Perancangan Aplikasi M-Discussion Berbasis Android Sebagai Wadah Diskusi Sekolah," *Syntax J. Inform.*, vol. 7, no. 2, p. 116, 2018.
- [5] A. Voutama and E. Novalia, "Perancangan Aplikasi M-Magazine Berbasis Android Sebagai Sarana Mading Sekolah Menengah Atas," *J. Tekno Kompak*, vol. 15, no. 1, pp. 104–115, Feb. 2021, Accessed: Oct. 28, 2022. [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/teknokompak/article/view/920>
- [6] S. D. Purnamasari and F. Panjaitan, "Pengembangan Aplikasi E-Reporting Kerusakan Lampu Jalan berbasis Mobile," *Jusikom J. Sist. Komput. Musirawas*, vol. 5, no. 1, pp. 59–69, 2020, doi: 10.32767/jusikom.v5i1.764.
- [7] D. Adhe, C. Rachman, R. Goejantoro, and D. Tisna, "Implementation Of Text Mining For Grouping Thesis Documents Using K-Means Clustering," *J. EKSPONENSIAL*, vol. 11, no. 2, pp. 167–174, 2020.
- [8] A. Yoga Pratama *et al.*, "Analisis Sentimen Media Sosial Twitter Dengan Algoritma K-Nearest Neighbor Dan Seleksi Fitur Chi-Square (Kasus Omnibus Law Cipta Kerja)," *J. Sains Komput.*

- Inform. (J-SAKTI)*, vol. 5, no. 2, pp. 897–910, 2021.
- [9] Euis Saraswati, Yuyun Umaidah, and Apriade Voutama, “Penerapan Algoritma Artificial Neural Network untuk Klasifikasi Opini Publik Terhadap Covid-19,” *Gener. J.*, vol. 5, no. 2, pp. 109–118, 2021, doi: 10.29407/gj.v5i2.16125.
- [10] Raksaka Indra Alhaqq, I Made Kurniawan Putra, and Yova Ruldeviyani, “Analisis Sentimen terhadap Penggunaan Aplikasi MySAPK BKN di Google Play Store,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 11, no. 2, pp. 105–113, 2022, doi: 10.22146/jnteti.v11i2.3528.
- [11] R. T. Aldisa and P. Maulana, “Analisis Sentimen Opini Masyarakat Terhadap Vaksinasi Booster COVID-19 Dengan Perbandingan Metode Naive Bayes, Decision Tree dan SVM,” *Technol. Sci.*, vol. 4, no. 1, pp. 106–109, 2022, doi: 10.47065/bits.v4i1.1581.
- [12] U. Kusnia and F. Kurniawan, “Analisis Sentimen Review Aplikasi Media Berita Online Pada Google Play menggunakan Metode Algoritma Support Vector Machines (SVM) Dan Naive Bayes,” *Explor. IT*, vol. 14, no. 36, pp. 24–28, 2022.
- [13] R. Akbar, Bagus Muhammad; Akbar, Ahmad Taufiq; Husaini, “Analisis Sentimen dan Emosi Vaksin Sinovac pada Twitter Menggunakan Naive Bayes dan Valence Shifter,” *J. Teknol. Terpadu*, vol. 7, no. 2, pp. 83–92, 2021, doi: <https://doi.org/10.54914/jtt.v7i2.433>.
- [14] D. P. Wilandini, “Penerapan Algoritma Naive Bayes Dalam Mengklasifikasikan Media Sosial Untuk Mengamati Trend Kuliner,” *J. Teknol. Terpadu*, vol. 8, no. 1, pp. 31–39, 2022, doi: <https://doi.org/10.54914/jtt.v8i1.535>.
- [15] M. F. Andriansyah, D. Yusup, and A. Voutama, “Menggunakan Metode Naive Bayes Berbasis Website Web-Based Expert System of Covid-19 Early Detection Using Naive Bayes Method,” *J. Inf. Technol. Comput. Sci.*, vol. 4, no. 2, pp. 446–455, 2021.
- [16] A. Voutama, “Sistem Antrian Cuci Mobil Berbasis Website Menggunakan Konsep CRM dan Penerapan UML,” *Komputika J. Sist. Komput.*, vol. 11, no. 1, pp. 102–111, 2022, doi: 10.34010/komputika.v11i1.4677.



PENGEMBANGAN APLIKASI *EMOTION RECOGNITION* DAN *FACIAL RECOGNITION* MENGGUNAKAN ALGORITMA *LOCAL BINARY PATTERN HISTOGRAM (LBPH)* DAN *CONVOLUTIONAL NEURAL NETWORK (CNN)*

Haeruddin ¹, Herman ², Patrick Pratama Hendri ³

^{1,2,3}Teknologi Informasi, Universitas Internasional Batam
Kota Batam, Kepulauan Riau, Indonesia 29426
haeruddin@uib.ac.id, herman@uib.ac.id, 1932024.patrick@uib.edu

Abstract

In the current modern era facial recognition technology can be found inside of everyday life, but said technology still has a big problem which is deepfake, where in which a deepfake can bypass security systems created with facial recognition as its base, one facial aspect that a deepfake cannot replicate perfectly is the emotion that can be observed from expression, which is why an emotion can be used as a tool to detect a deepfake, which is why an application that can detect both face and emotion at the same time is needed to add security to facial recognition technology, writer has succeeded in creating an application that can do both emotion recognition and facial recognition at the same time using LBPH (Local Binary Pattern Histogram) algorithm and purposive sampling technique for the facial recognition aspect with 67.5% accuracy and CNN (Convolutional Neural Network) algorithm using FER2013 (Facial emotion Recognition 2013) dataset for the emotion recognition aspect with 58.4% accuracy, with CRISP-DM method that can achieve the average accuracy rate of 63%, because currently not many research combine facial recognition using LBPH (Local Binary Pattern Histogram) algorithm and emotion recognition using CNN (Convolutional Neural Network) algorithm at the same time.

Keywords: AI, CNN, CRISP-DM, LBPH, Machine Learning

Abstrak

Pada zaman modern sekarang teknologi *facial recognition* sudah dapat ditemukan dalam kehidupan sehari-hari, akan tetapi teknologi ini masih memiliki satu masalah besar yaitu *deepfake*, dimana sebuah *deepfake* dapat mengelabui sistem keamanan yang dikembangkan berbasis *facial recognition*, salah satu aspek wajah yang belum dapat di replikasi dengan sempurna oleh *deepfake* adalah emosi yang dapat dilihat berdasarkan ekspresi wajah, oleh karena itu emosi dapat digunakan sebagai sebuah alat untuk mendeteksi sebuah *deepfake*, oleh karena itu dibutuhkan sebuah aplikasi yang dapat mendeteksi wajah dan emosi secara bersamaan untuk menambahkan keamanan teknologi *facial recognition*, penulis berhasil mengembangkan sebuah aplikasi yang dapat melakukan *emotion recognition* dan *facial recognition* secara bersamaan menggunakan algoritma LBPH (*Local Binary Pattern Histogram*) dengan data yang didapatkan menggunakan teknik *purposive sampling* untuk aspek *facial recognition* dengan keakuratan 67.5% dan algoritma CNN (*Convolutional Neural Network*) dengan menggunakan *dataset FER2013 (Facial emotion Recognition 2013)* untuk aspek *emotion recognition* dengan keakuratan 58.4%, menggunakan metode CRISP-DM yang dapat menghasilkan rata-rata keakuratan sebesar 63%, diakarenakan belum banyak penelitian yang menggabungkan *facial recognition* menggunakan algoritma LBPH (*Local Binary Pattern Histogram*) dengan *emotion recognition* menggunakan algoritma CNN (*Convolutional Neural Network*) secara bersamaan.

Kata kunci: AI, CNN, CRISP-DM, LBPH, Machine Learning

1. PENDAHULUAN

Pada zaman modern sekarang teknologi *facial recognition* sudah banyak dapat ditemukan dalam kehidupan sehari-hari manusia. Sekarang ini sebuah mesin dapat secara otomatis melakukan verifikasi informasi identitas dalam yang dapat digunakan kebutuhan keamanan dan pengawasan ataupun untuk mengakses kontrol ke dalam sebuah gedung [1].

Seiring dengan perkembangannya teknologi *facial recognition* bertambah juga kegunaannya dalam berbagai aspek kehidupan manusia, dari sisi keamanan teknologi ini dapat digunakan sebagai alat untuk mengidentifikasi kriminal menggunakan CCTV [2]. Dari sisi bisnis, teknologi ini memberikan sebuah pilihan baru sebagai alat pembayaran yang lebih aman [3]. Dari sisi edukasi juga

teknologi ini dapat mempermudah dan mempercepat proses absensi siswa dalam kelas [4]. Selain dari aspek-aspek tersebut teknologi *facial recognition* juga dapat ditemukan dalam *smartphone* modern sebagai alat autentikasi pengguna secara langsung [1]. Akan tetapi teknologi ini masih dapat menghasilkan dampak negatif yang ditimbulkan oleh segelintir manusia yang menyalahgunakan teknologi ini [5], salah satu diantaranya adalah pembuatan *deepfake*.

Deepfake adalah *video hyper-realistic* yang menggambarkan orang-orang mengatakan ataupun melakukan sesuatu yang tidak pernah terjadi dalam kenyataan [6]. *Deepfake* biasanya dibuat dengan memasukkan wajah seseorang dan sebuah *video* yang tidak ada hubungannya dengan orang tersebut kedalam sebuah algoritma *deep-learning* agar bisa didapatkan sebuah AI yang dapat mengubah wajah yang muncul dalam *video* tersebut menjadi wajah orang yang lain [6]. Sebuah *deepfake* berdasarkan kualitas produksinya dapat dibagi menjadi dua yaitu *low quality* dan *high quality*, sebuah *deepfake high quality* dapat menipu sistem autentikasi yang menggunakan *facial recognition* dengan tingkat keakuratan sebesar 85.62% sampai dengan 95% [7]. Oleh karena itu dibutuhkan sebuah teknologi yang dapat mendeteksi sebuah *deepfake*.

Salah satu cara yang dapat digunakan untuk mendeteksi sebuah *deepfake* adalah dengan menggunakan pendeteksi emosi atau *emotion recognition*. Dikarenakan sebuah *deepfake* yang diproduksi dengan menggunakan algoritma *face-swap* belum dapat menciptakan kembali ekspresi yang sesuai dengan emosi [8]. Kemudian lagi sebuah *deepfake* juga belum bisa menghasilkan ekspresi emosi yang konsisten dengan emosi yang dideteksi dari dalam suaranya [9]. Oleh karena itu pengenalan emosi atau *emotion recognition* dapat digunakan sebagai lapisan tambahan keamanan dalam sebuah aplikasi *facial recognition* untuk mengurangi dampak negatif dari *deepfake*.

Pengembangan teknologi *facial recognition* dapat dilakukan menggunakan bahasa pemrograman *python* dengan *library opencv* (*Open Computer Vision*) yang dapat mempermudah dan memberikan peralatan – peralatan yang dapat digunakan dalam proses pengembangan aplikasi *facial recognition* [10]. Salah satu algoritma yang digunakan dalam pembuatan aplikasi *facial recognition* adalah algoritma LBPH (*Local Binary Pattern Histogram*) yang direkomendasikan oleh [11] dikarenakan aplikasi yang dihasilkan dapat memberikan hasil yang bagus dari berbagai eksperimen dalam skenario yang telah ditentukan. Kemudian untuk pengembangan aplikasi *emotion recognition* yang dapat mendeteksi emosi yang sedang dirasakan oleh subjek yang terdeteksi secara real-time dengan menggunakan algoritma CNN (*Convolutional Neural Network*), algoritma tersebut dapat menghasilkan sebuah aplikasi yang dapat mencapai tingkat keakuratan sebesar 89.98% [12].

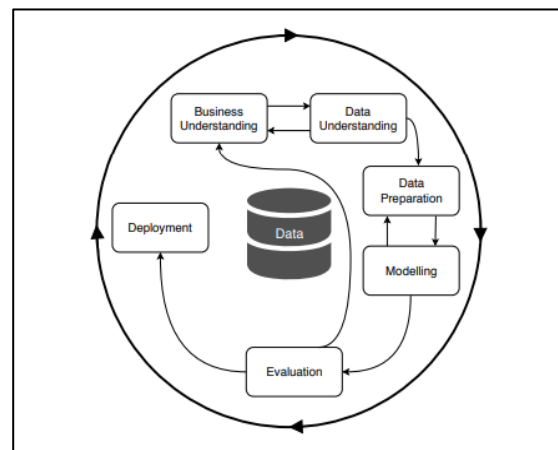
Penelitian ini dilakukan untuk menghasilkan sebuah aplikasi yang dapat mendeteksi wajah dan emosi subjek secara *real-time* yang menggunakan sampel wajah ras asia tenggara. Penelitian ini dilakukan dengan metode CRISP-DM (*Cross-industry standard process for data mining*) sebagai kerangka dalam pengembangan aplikasi *emotion recognition* dan *facial recognition* menggunakan algoritma LBPH (*Local Binary Pattern Histogram*) dan CNN (*Convolutional Neural Network*). Diakarenakan belum banyak penelitian yang menggabungkan *facial recognition* menggunakan algoritma LBPH (*Local Binary Pattern Histogram*) dengan *emotion recognition* menggunakan algoritma CNN (*Convolutional Neural Network*) secara bersamaan. Penulis berharap dengan adanya penelitian ini dapat membantu penelitian lainnya yang membutuhkan aplikasi yang bisa merekognisi wajah dan emosi secara bersamaan di masa depan.

2. METODE PENELITIAN

Dalam perancangan aplikasi *emotion recognition* dan *facial recognition* menggunakan algoritma LBPH (*Local Binary Pattern Histogram*) dan CNN (*Convolutional Neural Network*) penulis menggunakan metode CRISP-DM seperti [13] sebagai kerangka proses pengembangan aplikasi. Penelitian ini juga menggunakan *dataset* FER2013 [12] sebagai *dataset* pendukung pembuatan model *emotion recognition*. Berikut penjelasan lebih lanjut metode penelitian yang dilakukan oleh penulis.

2.1 Metode Pengembangan

Pada penelitian “Pengembangan Aplikasi *Emotion Recognition* dan *Facial Recognition* Menggunakan Algoritma LBPH (*Local Binary Pattern Histogram*) dan CNN (*Convolutional Neural Network*)” penulis menggunakan metode CRISP-DM (*Cross-industry standard process for data mining*) sebagai kerangka pengembangan aplikasi. Dikarenakan metode CRISP-DM masih digunakan sebagai kerangka umum dalam proses pengelolaan data [13]. Dimana tahapannya ditunjukkan oleh gambar 1.



Gambar 1. Proses CRISP-DM [13]

a. Business Understanding

Pada tahap ini penulis melakukan studi literatur dan mendapatkan bahwa *deepfake* sebagai salah satu kelemahan dari teknologi *facial recognition* dan menemukan cara untuk mengatasi masalah tersebut. Berdasarkan [8], [9] emosi dapat digunakan sebagai sebuah cara untuk mendeteksi *deepfake* mulai dari kualitas yang rendah sampai dengan kualitas yang tinggi.

b. Data Understanding

Pada tahap ini penulis akan mengumpulkan dan menganalisis kualitas data yang akan digunakan. Data yang digunakan untuk melatih *facial recognition* merupakan 100 gambar wajah dan nama untuk 5 subjek yang dipilih dengan menggunakan teknik *purposive sampling* dan *dataset FER2013 (Facial Expression Recognition 2013)* sebagai data landasan model *emotion recognition*.

c. Data Preparation

Pada tahap ini penulis akan memulai mempersiapkan data yang didapatkan dari tahap *data preparation* untuk proses selanjutnya. Penulis melakukan pembersihan data atau *data cleaning* ke dalam data wajah yang telah dikumpulkan, dengan cara menghapus gambar wajah yang tidak *valid* dengan *parameter* sebuah wajah wajib terlihat jelas dan lengkap memiliki 2 mata, sebuah hidung, dan mulut yang terlihat dengan jelas seperti yang terlihat pada gambar 2.



Gambar 2. Contoh gambar wajah yang valid, dimana terlihat 2 mata, sebuah hidung, dan mulut yang jelas

Kemudian penulis akan mengubah *dataset FER2013* dari bentuk *csv* menjadi foto-foto yang dikelompokkan ke dalam *folder* tersendiri berdasarkan jenis emosi dan kegunaan foto sebagai *data training* atau *data testing*.

d. Modelling

Pada tahap ini penulis akan mulai membuat *model* yang akan digunakan dalam aplikasi yang akan dihasilkan menggunakan data yang telah disiapkan pada tahap *data preparation*. Untuk membangun model aspek *facial recognition*, akan digunakan algoritma LBPH yang akan mengasosiasikan aspek wajah yang ada dengan nama subjek yang sesuai. Kemudian penulis menggunakan algoritma CNN dengan 28709 data emosi yang digunakan dan dilatih

selama 50 *epoch* untuk menghasilkan model yang akan digunakan dalam aspek *emotion recognition*.

Algoritma yang digunakan dalam penelitian ini adalah, sebagai berikut:

- LBPH (*Local Binary Pattern Histogram*)
LBPH (*Local Binary pattern Histogram*) sejak penemuannya merupakan algoritma yang cukup bagus untuk mengklasifikasikan tekstur-tekstur tertentu seperti wajah, algoritma ini membutuhkan 4 *parameter* unik untuk memproses sebuah gambar, yaitu : *radius* (r), *neighbors* (n), *X-axis* dan *Y-axis* [14].
- CNN (*Convolutional Neural Network*)
Metode CNN dihasilkan dari pengembangan metode *Multilayer Perceptron (MLP)* untuk memproses data dalam 2 dimensi [15]. Cara kerja CNN hampir sama dengan MLP akan tetapi, dengan MLP setiap *neuron* hanya terdapat dalam 1 dimensi sedangkan pada CNN setiap *neuron* terdiri dari 2 dimensi [16].

e. Evaluation

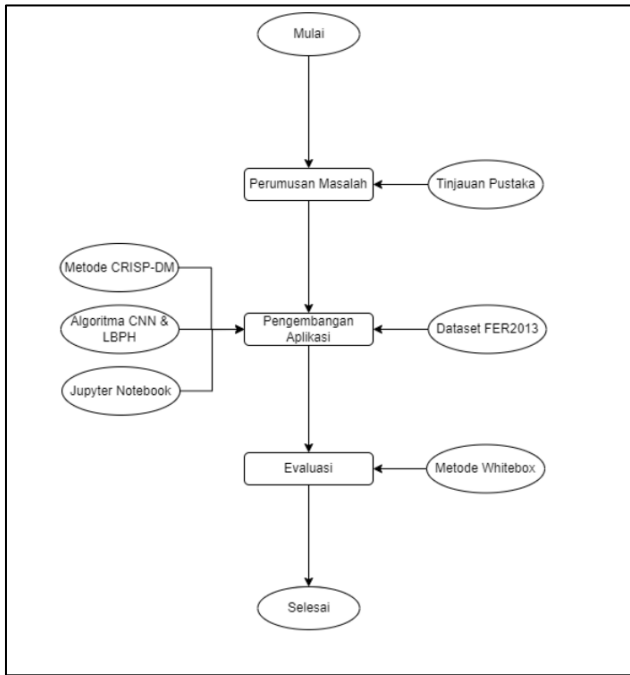
Pada tahap ini penulis mengevaluasi *model* yang telah dihasilkan dari tahap *modelling* dengan metode *software testing whitebox*, dimana tingkat akurasi dapat dihasilkan dengan mengevaluasi logika internal dan struktur program dengan melakukan evaluasi berdasarkan hasil dari skenario yang telah ditentukan dan hasil yang seharusnya didapatkan [17].

f. Deployment

Tahap *deployment* tidak dilakukan pada penelitian ini, dikarenakan penelitian ini bersifat *experimental* studi yang mana bertujuan untuk menguji penerapan algoritma LBPH & CNN dan belum *di-deploy* kedalam aplikasi tertentu.

2.2 Tahapan penelitian

Berikut terlampir tahapan penulis dalam melakukan penelitian “Pengembangan Aplikasi *Emotion Recognition* dan *Facial Recognition* Menggunakan Algoritma LBPH (*Local Binary Pattern Histogram*) dan CNN (*Convolutional Neural Network*)” yang ditunjukkan pada gambar 3.



Gambar 3 Research Flow

Berikut merupakan penjelasan dari alur penelitian diatas :

1. Pada tahap pertama penulis mulai dari perumusan masalah berdasarkan topik yang diangkat untuk menentukan arah dan hasil dari penelitian ini.
2. Pada tahap kedua penulis mulai mengembangkan aplikasi *facial recognition* dengan algoritma LPBH dan *emotion recognition* dengan algoritma CNN menggunakan metode CRISP-DM sebagai kerangka proses pengembangan aplikasi. Penulis juga menggunakan *dataset FER2013* sebagai data yang digunakan untuk membuat model *emotion recognition*.
3. Pada tahap ketiga penulis akan mengevaluasi performa dari aplikasi yang telah selesai dikembangkan dengan metode *software testing whitebox* untuk mendapatkan tingkat keakuratan aplikasi.
4. Pada tahap terakhir setelah penulis mendapatkan hasil performa aplikasi, penulis akan menarik kesimpulan sesuai dengan hasil yang didapatkan serta memberikan rekomendasi kepada pembaca.

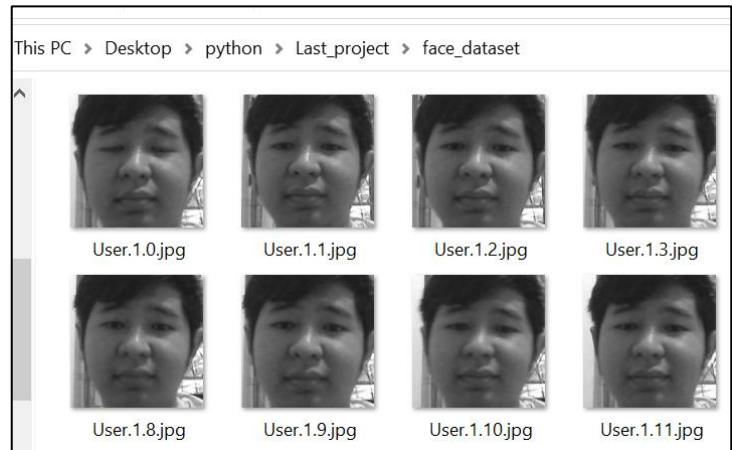
3. HASIL DAN PEMBAHASAN

Hasil yang didapatkan oleh penelitian ini menunjukkan bahwa algoritma LBPH dapat mendeteksi fitur wajah secara *real-time* dengan hasil keakuratan sebesar 67.5% yang didukung oleh penelitian yang dilakukan oleh Farah Deeba, et.al pada tahun 2019 [11]. Begitu juga dengan algoritma CNN yang dapat mendeteksi emosi secara *real-time* dengan hasil keakuratan 58.4 % dan didukung oleh penelitian oleh Rohit Pathar, et.al pada tahun 2019 [12]. Selain itu kedua algoritma ini juga dapat mendapatkan rata-rata 63 % tingkat keakuratan ketika digunakan secara bersamaan. Hasil evaluasi dapat dilihat dalam tabel 1.

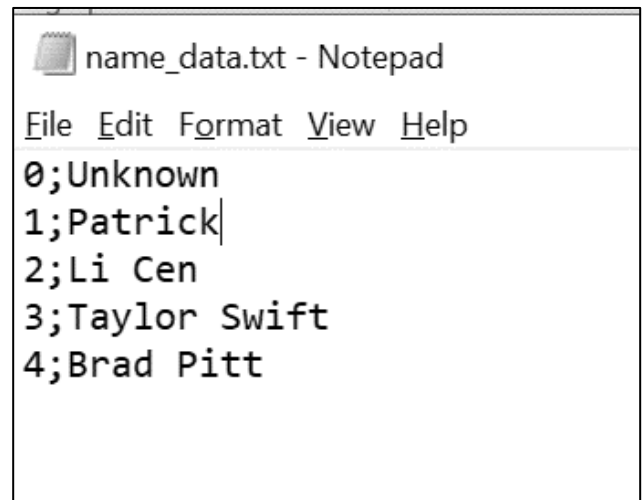
Tabel 1. Tabel Hasil Evaluasi Algoritma

Test dalam real-time	Kebenaran		Total Data	Total (%)
	Ya	Tidak		
Facial Recognition	52	25	77	67.5
Emotion Recognition	45	32	77	58.4
Rata-rata				63

a. Facial Recognition

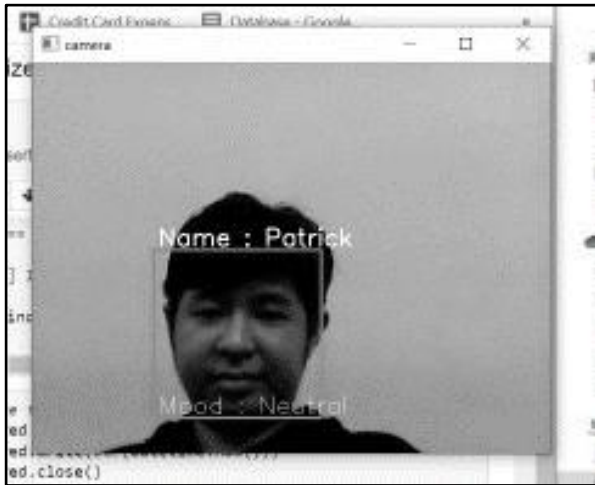


Gambar 4. Data Wajah yang Didapatkan akan Disimpan Per User sebanyak 100 kali



Gambar 5. Penyimpanan Nama dan ID Subjek

Sebelum dapat digunakan untuk mendeteksi wajah secara *real-time* algoritma LBPH membutuhkan *dataset* wajah yang akan diasosiasikan dengan nama subjek tersebut. Dapat dilihat dari gambar 4 dan 5 penulis menyimpan data wajah subjek dan nama subjek yang sesuai dengan nama *file* yang telah didapatkan.



Gambar 6. LBPH Berhasil Merekognisi Wajah, dimana identitas subjek dapat direkognisi dengan benar

Berdasarkan gambar 6 dapat dilihat bahwa algoritma LBPH dapat merekognisi wajah subjek apabila *dataset* wajah subjek telah diregistrasikan kedalam algoritma. Penulis juga menguji algoritma ini lebih lanjut dengan menguji lebih dari satu subjek secara *real-time*.



Gambar 7. Hasil Rekognisi Dua Wajah Sekaligus, dimana wajah dan emosi subjek 1 (kiri) dapat diidentifikasi dengan benar sedangkan wajah subjek 2 (kanan) tidak dapat di-identifikasi dengan benar



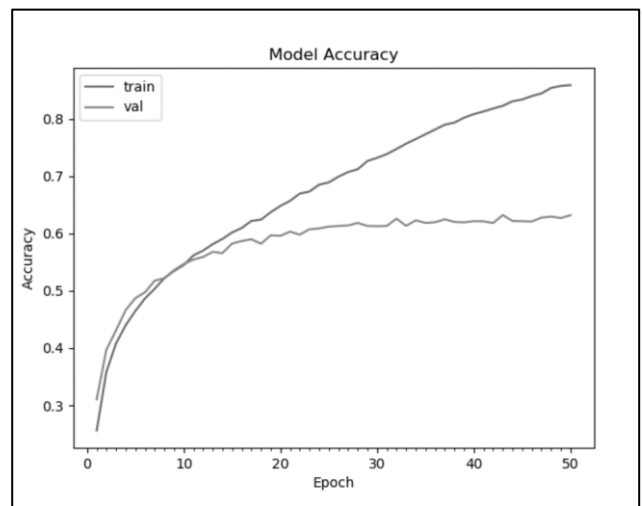
Gambar 8. Hasil Rekognisi Tiga Wajah sekaligus, dimana 2 subjek bagian bawah dapat teridentifikasi dengan benar sedangkan 1 subjek bagian atas tidak dapat diidentifikasi dengan benar

Akan tetapi algoritma LBPH masih memiliki kelemahan dalam merekognisi wajah, apabila subjek yang didapatkan lebih dari satu orang. Dapat dilihat dari gambar 7 dimana dari dua wajah yang didapatkan hanya satu yang nama dari subjek tersebut dapat ditemukan, dan gambar 8 dimana dari tiga subjek hanya dua subjek yang dapat terdeteksi.

Secara keseluruhan algoritma LBPH dapat menghasilkan rata-rata tingkat keakuratan sebesar 67.5% dari 77 kali *testing* dimana 40 diantaranya merupakan satu subjek secara *real-time* dan 37 sisanya merupakan lebih dari satu subjek secara langsung.

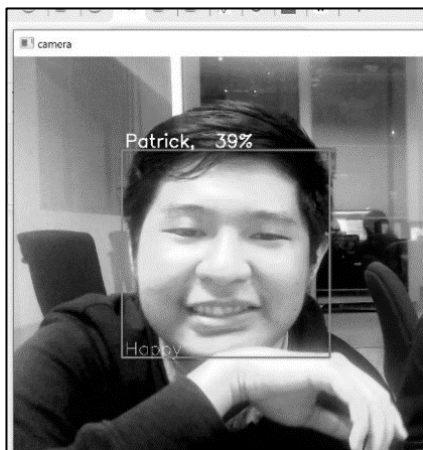
b. Emotion Recognition

Tingkat keakuratan dari model yang dihasilkan menggunakan algoritma CNN dapat diperhatikan dalam gambar 9. Dimana dari *dataset* FER2013 yang berisi dengan 35887 dipisahkan menjadi 28709 *data training* dan 7178 *data validasi*, yang dilakukan selama 50 *epoch*.



Gambar 9. Kurva Keakuratan *Data Training* vs *Data Validasi* selama 50 *Epoch*

Untuk mengevaluasi ketepatan algoritma secara *real-time* penulis melakukan *testing* sebanyak 77 kali, dimana 40 diantaranya merupakan satu subjek secara *real-time* dan 37 sisanya merupakan lebih dari satu subjek secara langsung. Hasil yang didapatkan menunjukkan akurasi yang cukup baik apabila subjek yang dideteksi dapat menunjukkan emosi secara jelas seperti yang terlihat dalam gambar 10.



Gambar 10. Aplikasi berhasil mendeteksi emosi dengan benar dimana dapat terlihat subjek sedang senang

Tetapi apabila posisi ataupun emosi subjek tidak dapat terlihat dengan jelas, maka aplikasi akan memberikan hasil prediksi emosi yang dianggap paling dekat dengan atribut wajah subjek yang sedang dideteksi seperti yang terlihat pada gambar 11.



Gambar 11. Aplikasi memberikan prediksi emosi yang tidak tepat dimana wajah subjek terlihat tidak marah

Secara keseluruhan algoritma CNN dapat menghasilkan rata-rata tingkat keakuratan 58.4 % dalam mendeteksi emosi wajah secara *real-time*.

4. KESIMPULAN

Berdasarkan hasil evaluasi yang telah dilakukan, aplikasi yang dapat melakukan *facial recognition* menggunakan algoritma LBPH dan *emotion recognition* dengan algoritma CNN berhasil dihasilkan. Berikut deskripsi proses aplikasi yang telah dihasilkan:

1. Proses pembuatan model menggunakan algoritma LBPH untuk aspek *facial recognition* dan algoritma CNN dengan *dataset FER2013* untuk aspek *emotion recognition*, dimana kedua algoritma ini dapat melakukan proses ekstraksi fitur dan prediksi wajah.
2. Rekognisi wajah dan emosi secara *real-time* menggunakan LBPH untuk klasifikasi wajah seakurat

67.5% dan CNN untuk klasifikasi emosi seakurat 58.4%.

3. Ketika aplikasi sedang berjalan, informasi klasifikasi wajah dan emosi yang dideteksi akan dimunculkan secara langsung di atas tampilan kamera yang diambil.

Dibutuhkan penelitian yang lebih lanjut untuk dapat menghasilkan model yang dapat membaca wajah dan emosi dari ekspresi dimana terdeteksi lebih dari 1 subjek secara efektif.

Ucapan Terima Kasih

Penulis P.P.H. mengucapkan terima kasih kepada Direktorat Universitas Internasional Batam, Bapak Dr. Hendi Sama selaku Dekan Fakultas Sistem Informasi Universitas Internasional Batam, orang tua dan keluarga besar penulis, dan sahabat dan teman-teman penulis.

DAFTAR PUSTAKA

- [1] M. I. Zarkasyi, M. R. Hidayatullah, and E. M. Zamzami, "Literature Review: Implementation of Facial Recognition in Society," *J. Phys. Conf. Ser.*, vol. 1566, no. 1, 2020, doi: 10.1088/1742-6596/1566/1/012069.
- [2] N. B. Aherwadi, D. Chokshi, S. Pande, and A. Khamparia, "Criminal Identification System using Facial Recognition," *SSRN Electron. J.*, 2021, doi: 10.2139/ssrn.3884827.
- [3] Y. Zhong and H. C. Moon, "Investigating Customer Behavior of Using Contactless Payment in China: A Comparative Study of Facial Recognition Payment and Mobile QR-Code Payment," *Sustain.*, vol. 14, no. 12, 2022, doi: 10.3390/su14127150.
- [4] S. Bussa, A. Mani, S. Bharuka, and S. Kaushik, "Smart Attendance System using OPENCV based on Facial Recognition," *Int. J. Eng. Res.*, vol. V9, no. 03, pp. 54–59, 2020, doi: 10.17577/ijertv9is030122.
- [5] I. Arifin, R. F. Haidi, and M. Dzalhaqi, "Penerapan Computer Vision Menggunakan Metode Deep Learning pada Perspektif Generasi Ulul Albab," *J. Teknol. Terpadu*, vol. 7, no. 2, pp. 98–107, 2021, doi: 10.54914/jtt.v7i2.436.
- [6] M. Westerlund, "The emergence of deepfake technology: A review," *Technol. Innov. Manag. Rev.*, vol. 9, no. 11, pp. 39–52, 2019, doi: 10.22215/TIMREVIEW/1282.
- [7] P. Korshunov and S. Marcel, "Vulnerability assessment and detection of Deepfake videos," *2019 Int. Conf. Biometrics, ICB 2019*, 2019, doi: 10.1109/ICB45273.2019.8987375.
- [8] J. M. López-Gil, R. Gil, and R. García, "Do Deepfakes Adequately Display Emotions? A Study on Deepfake Facial Emotion Expression," *Comput.*

- Intell. Neurosci.*, vol. 2022, p. 1332122, 2022, doi: 10.1155/2022/1332122.
- [9] B. Hosler *et al.*, “Do deepfakes feel emotions? A semantic approach to detecting deepfakes via emotional inconsistencies,” *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 1013–1022, 2021, doi: 10.1109/CVPRW53098.2021.00112.
- [10] T. Dhawle, U. Ukey, and R. Choudante, “IRJET-Face Detection and Recognition using OpenCV and Python,” *Int. Res. J. Eng. Technol.*, vol. 7, no. 10, 2020.
- [11] F. Deeba, A. Ahmed, H. Memon, F. A. Dharejo, and A. Ghaffar, “LBPH-based enhanced real-time face recognition,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 5, pp. 274–280, 2019, doi: 10.14569/ijacsa.2019.0100535.
- [12] R. Pathar, A. Adivarekar, A. Mishra, and A. Deshmukh, “Human Emotion Recognition using Convolutional Neural Network in Real Time,” *Proc. 1st Int. Conf. Innov. Inf. Commun. Technol. ICICT 2019*, pp. 1–7, 2019, doi: 10.1109/ICICT1.2019.8741491.
- [13] F. Martinez-Plumed *et al.*, “CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories,” *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 8, pp. 3048–3061, 2021, doi: 10.1109/TKDE.2019.2962680.
- [14] L. Wang and A. A. Siddique, “Facial recognition system using LBPH face recognizer for anti-theft and surveillance application based on drone technology,” *Meas. Control (United Kingdom)*, vol. 53, no. 7–8, pp. 1070–1077, 2020, doi: 10.1177/0020294020932344.
- [15] J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, “Striving for simplicity: The all convolutional net,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Work. Track Proc.*, pp. 1–14, 2015.
- [16] L. Zahara, P. Musa, E. Prasetyo Wibowo, I. Karim, and S. Bahri Musa, “The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi,” *2020 5th Int. Conf. Informatics Comput. ICIC 2020*, vol. 7, 2020, doi: 10.1109/ICIC50835.2020.9288560.
- [17] S. Nidhra, “Black Box and White Box Testing Techniques - A Literature Review,” *Int. J. Embed. Syst. Appl.*, vol. 2, no. 2, pp. 29–50, 2012, doi: 10.5121/ijesa.2012.2204.



DESIGN SYSTEM PADA PERANCANGAN ANTAR MUKA PERANGKAT LUNAK SISTEM AKSES DIGITAL

Apriansyah Rizqi Setiawan¹, Marsani Asfi², Agus Sevtiana³, Sudadi Pranata⁴, Willy Eka Septian⁵

^{1,2,3,5}Fakultas Teknologi Informasi, Universitas Catur Insan Cendekia

⁴Fakultas Ekonomi dan Bisnis, Universitas Catur Insan Cendekia

Kota Cirebon, Jawa Barat, 45133

apriansyahrizs@gmail.com, marsani.asfi@cic.ac.id, agus.sevtiana@cic.ac.id, sudadi.pranata@cic.ac.id,
willy.eka.septian@cic.ac.id

Abstract

In developing software systems, there were no rules for making interface patterns in web base applications. The problem is the inconsistency of design styles between software, and the interface design development process takes a long time. This study aims to produce a design framework called Access Digital Design System. The system development method approach used in this research is Atomic Design which has several stages adapted to the investigation. In this study, the design system built is limited to product elements consisting of functional patterns and perceptual patterns. This research will produce a design system made up of perceptual and available marks as output artifacts in the form of code libraries, style guides, and pattern libraries. Based on the study, the designed Design System can reduce time spent creating user interface designs by using component guidelines and documentation with standardized foundation elements to increase productivity. Additionally, the final interface design becomes more consistent and uniform.

Keywords: Consistency, Design System, Efficiency, User Interface, User Experience

Abstrak

Dalam mengembangkan sistem perangkat lunak, selama ini belum ada aturan tentang pola standar *interface* pada perangkat lunak berbasis web. Masalah inkonsistensi pada gaya desain antar perangkat lunak dan waktu yang dibutuhkan dalam pengembangan desain *interface* cenderung menghabiskan waktu. Penelitian ini bertujuan untuk menghasilkan kerangka kerja desain bernama Akses Digital Design System. Pendekatan metode perancangan desain sistem yang digunakan dalam penelitian ini adalah *Atomic Design*. Pada penelitian ini *design system* yang dibangun dibatasi berupa elemen *product* yang terdiri dari *functional pattern* dan *perceptual pattern*. Adapun hasil atau artefak keluaran menghasilkan *design system* yang terdiri *perceptual pattern* dan *functional pattern* dalam bentuk *pattern library*, *style guide* dan *code library*. Dari hasil penelitian, *design system* dapat meningkatkan produktivitas dalam mendesain antarmuka pengguna sehingga waktu yang diperlukan dapat dikurangi. Hal ini dikarenakan adanya dokumentasi sebagai panduan terkait komponen-komponen yang akan digunakan sebagai elemen dasar standar. Selain itu, desain *user interface* yang dihasilkan lebih menyatu dan konsisten.

Kata kunci: Design System, Efisiensi, Konsistensi, User Experience, User Interface

1. PENDAHULUAN

Pada saat mengembangkan sistem perangkat lunak tidak ada aturan untuk membuat pola *user interface* pada sistem. Oleh karena itu *designer* dalam pembuatan *user interface* dari sistem perangkat lunak selama ini berdasarkan eksplorasi. Masalah dalam gaya yang berbeda dalam mendesain antar muka sistem yang dibangun oleh *developer* menyebabkan inkonsistensi tampilan *user interface*. Hal ini tentu saja akan mempengaruhi *user experience*. Ketika dalam proses perancangan *user interface*, *designer* cenderung terlalu fokus pada detail visual, yang membutuhkan waktu. Selain

itu, hasil *user interface* yang dibuat tidak konsisten selama *development*, *developer* juga melakukan upaya yang berlebihan karena ketidakkonsistenan dalam desain. Desain harus memiliki konsep atau jenis informasi yang sama, tetapi menggunakan pendekatan yang berbeda.

Berdasarkan permasalahan tersebut, maka diperlukan suatu *framework* desain agar adanya dokumentasi dan *design language*. Salah metodenya adalah *design system*. “*Design system* merupakan seperangkat pola yang saling berhubungan serta langkah sistematis dan kolaboratif yang

diatur secara koheren untuk mempercepat dan mempermudah proses desain dan *development* dalam mencapai tujuan produk digital” [1]. Dalam perspektif pengalaman pengguna adanya konsistensi merupakan salah satu faktor penting yang berpengaruh terhadap kemudahan dalam menggunakan desain. Konsistensi dalam mendesain menghindari pengguna dari kebingungan serta diperolehnya nilai baik pada kepuasan pengguna. Oleh karena itu, *design system* berperan penting bagi *startup* khususnya di Indonesia [2].

Beberapa penelitian tentang *interface* perangkat lunak yaitu digunakan untuk mengevaluasi kualitas suatu *website*. Prinsip *heuristic evaluation* digunakan untuk evaluasi. Dari hasil penelitian diperoleh lima prinsip yang memenuhi aturan *heuristic evaluation* yaitu visibilitas status sistem, kesesuaian antara sistem dengan dunia nyata, standar dan konsisten, memahami lebih baik dari mengingat, *flexibilitas* dan efisien. Sedangkan lima prinsip yang tidak terpenuhi yaitu kendali dan kebebasan pengguna, estetika, pencegahan kesalahan, desain secara minimalis, bantuan untuk pengguna dalam mengenali, berdialog serta memperbaiki kesalahan, serta adanya dokumentasi. Pengujian terhadap performa *website* menunjukkan nilai 75, hal ini berarti sudah cukup baik [3].

Penelitian lainnya [4] tentang UI/UX pada *online shop*. Hasil penelitian menunjukkan pengaruh dari *User Interface* (UI) dan *User Experience* (UX) dalam merancang desain *prototype* suatu *website*. Proses perancangan yang digunakan adalah metode *Design Thinking* dengan *emphatize, define, ideate, prototype dan test* [4]. Penelitian lainnya terkait dengan perancangan *web mobile prototype* desain untuk sistem jual beli. Desain menggunakan Figma. Penelitian lainnya [13] tentang rancang bangun sistem informasi agroindustri. Hasil dari desain antarmuka memudahkan pengembangan sistem selanjutnya dan kemudahan bagi pengguna terhadap tampilan yang *user friendly*. Hasil penelitian berupa desain tampilan *web mobile* [5]. Begitu juga penelitian yang dilakukan [6] penggunaan metode *design thinking* serta pengujian dengan *system usability Scale*. Sedangkan penelitian yang dilakukan [7] penggunaan metode *atomic design* dalam desain sistem di suatu *website* perguruan tinggi.

Penelitian lainnya dilakukan [8] adalah menggunakan metode *User Centered Design* (UCD). Dalam penelitian tersebut dilakukan sketsa perancangan antarmuka aplikasi serta komponen-komponen antarmuka. Perancangan sketsa menggunakan sistem desain *material design* sedangkan untuk merancang antarmuka dengan *high fidelity* dilakukan menggunakan perangkat lunak Figma [8].

Penelitian terkait desain *user interface* untuk aplikasi *virtual tour* pengenalan lingkungan kampus dirancang dengan aplikasi *balsamiq* [9]. Hasil penelitian menghasilkan tampilan menjadi lebih menarik dan interaktif [9], penelitian lainnya terkait *design system* adalah melalui *design system*

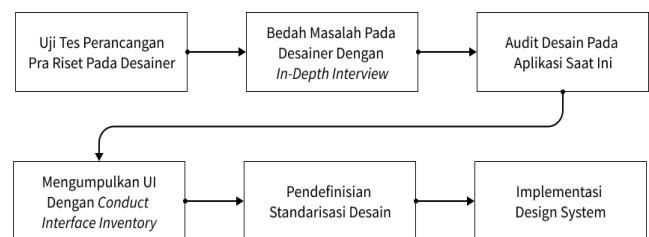
untuk mempermudah masinis untuk mengetahui kondisi kereta api [10].

Dari uraian di atas permasalahan dalam penelitian ini yaitu belum adanya aturan dalam menerapkan pola *user interface* pada sistem, *designer user interface* sistem perangkat lunak masih berdasarkan eksplorasi. Selain itu permasalahan dalam gaya yang berbeda dalam mendesain antar muka sistem yang menyebabkan inkonsistensi tampilan *user interface, designer* juga cenderung terlalu fokus pada detail visual, yang membutuhkan waktu.

Tujuan dari penelitian ini adalah bagaimana menerapkan konsep atau jenis informasi yang sama dalam proses perancangan aplikasi berdasarkan pendekatan yang berbeda. Manfaat yang akan diperoleh dari penelitian ini adalah: (1) adanya konsistensi dalam perancangan *user interface*, (2) adanya standar yang seragam dalam mendesain antar muka sistem, (3) memiliki panduan utama dalam merancang desain antar muka sistem, karena objek-objek yang digunakan memiliki konsistensi serta keseragaman dalam merancang *interface* sistem. Metode yang digunakan dalam perancangan antarmuka adalah *design system* dengan pendekatan *atomic system*. *Design system* digunakan dalam penelitian ini untuk menyelesaikan permasalahan dalam perancangan desain tampilan karena menerapkan pola yang saling berhubungan dengan langkah-langkah yang sistematis dan kolaboratif. Langkah-langkah ini diatur secara koheren agar dapat mempercepat dan mempermudah proses desain dan *development* dalam mencapai tujuan produk digital.

2. METODE PENELITIAN

Metodologi penelitian yang dilakukan dalam penelitian ini sesuai dengan diagram penelitian seperti pada Gambar 1.



Gambar 1. Kerangka Penelitian

Gambar 1 kerangka penelitian yang dilakukan adalah :

2.1. Uji Test Perancangan Pra-riiset pada Desainer

Pada tahap ini pengujian pra-riiset dilakukan untuk mengetahui masalah apa yang ditemukan dan juga waktu yang dibutuhkan untuk menyelesaikan tugas pengujian. Tugas yang diberikan berupa *case study* yang mengharuskan desainer membuat tampilan *user interface*. Kemudian hasil dari pengujian ini akan dijadikan pembandingan sebagai acuan dalam pengujian *design system*.

2.2 Bedah Masalah pada Desainer dengan Metode *In-depth Interview*

Pada tahap ini dilakukan pengumpulan tentang data perilaku dan pengalaman desainer ketika merancang *user interface*. Teknik *in-depth interview* melibatkan wawancara intensif dengan responden untuk mengeksplorasi perspektif ide, program atau situasi tertentu [11].

2.3. Audit Desain Perangkat Lunak

Di tahapan ini dilakukan desain audit pada salah satu desain perangkat lunak yang telah dikembangkan di perusahaan. Audit dilakukan untuk mencatat dan mengumpulkan semua komponen *user interface* yang terdapat pada perangkat lunak. Selanjutnya dilakukan kategorisasi berdasarkan jenis komponennya sehingga diperoleh daftar-daftar komponen yang tidak konsisten [12].

2.4. Pendefinisian Standarisasi Desain

Pada tahap ini dilakukan pengumpulan untuk semua komponen yang ter-kategorisasi. Di tahapan ini juga ditentukan satu atau beberapa *style* yang dijadikan standar untuk setiap komponen. Selanjutnya ditentukan *pattern library* dan *style guideline* sebagai brand atau identitas dari desain serta dibuatkan panduan *study case* dalam menggunakan komponen desain [12].

2.5. Implementasi *Design System*

Pada tahapan ini dilakukan proses implementasi *design system* dengan melakukan pembentukan *pattern library* dari hasil perancangan model konseptual. Dalam tahapan ini akan dibangun juga dokumentasi berupa *framework* desain sebagai bentuk *deliverables* berupa dokumentasi panduan desain.

3. HASIL DAN PEMBAHASAN

3.1. Pengujian Pra Penelitian

Hasil dari pengujian pra-penelitian diperoleh lamanya waktu dalam membuat *user interface* oleh *designer*. Hasil pengujian dari *designer* dilakukan dengan skenario tugas yang dibagi dalam beberapa tugas untuk membuat satu halaman *website*.

Tabel 1. Pengujian Pra Penelitian

Deskripsi	Durasi
Tugas 1	26 menit 45 detik
Tugas 2	45 menit 10 detik
Tugas 3	48 menit 20 detik
Total	2 jam 15 detik
Rata-rata	40 menit 5 detik

Dari perolehan waktu pada Tabel 1 rata-rata waktu penyelesaian tugas pra-penelitian adalah 40 menit 5 detik. Data pra-penelitian ini selanjutnya dijadikan dasar perbandingan dengan pengujian pasca-penelitian.

3.2 *In-Depth Interview*

Hasil *In-Depth Interview* bertujuan untuk mengetahui masalah secara dalam tentang apa saja yang dialami oleh *designer*. Kesimpulan hasil wawancara seperti terlihat pada Tabel 2.

Tabel 2. Hasil *In-Depth Interview*

No.	Kesimpulan
1.	Desainer mengalami kebingungan akan konsep <i>user interface</i> .
2.	Pada <i>user interface</i> , <i>designer</i> elemen untuk setiap komponen dibuat dari awal sehingga membutuhkan waktu.

3.3. *Audit Interface*

Hasil *audit interface* seperti terlihat pada Gambar 2 merupakan komponen-komponen yang tidak konsisten. Komponen antarmuka yang tidak konsisten tersebut meliputi elemen-elemen umum. Tabel 3 adalah hasil audit *website* CV Akses Digital. Pada Tabel 3 tersebut terlihat beberapa jenis elemen *button* yang digunakan pada *website*.



Gambar 2. Hasil *Audit Interface*

3.4. *Interface Inventory*

Hasil dari *interface inventory* diperoleh melalui identifikasi perilaku dari pengguna, uraian aksi spesifik yang dihasilkan serta pengumpulan dari komponen. Hasil dari *interface inventory* menjadi pengalaman bagi pengguna ketika berinteraksi dengan *website*.

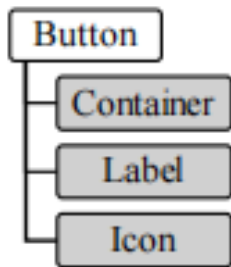
Tabel 3. Hasil *Interface Inventory*

Kategori	Komponen
<i>Basic Button</i>	
<i>Button Icon</i>	
<i>Button Link</i>	

3.5. Perancangan *Design System*

1. Struktur Komponen

Perancangan *design system* dilakukan dengan menguraikan struktur pada tiap komponen. Tiap suatu komponen terdiri dari bagian elemen elemen terkecil sesuai dengan *Atomic Design Methodology*.



Gambar 3. Struktur Komponen Perancangan *Design System*

Pada Gambar 3 merupakan komponen *button* yaitu *container*, *label* serta *icon*. Komponen-komponen *button* merupakan bentuk lain dari komponen *button*. Varian komponen *button* terlihat pada Tabel 4.

Tabel 4. Varian Komponen Button

Varian	Keterangan
<i>Basic Button</i>	Digunakan sebagai suatu tindakan pada antarmuka. Contoh : berpindah antar halaman atau melakukan <i>submit</i> pada suatu <i>form</i> .
<i>Button dengan icon</i>	Digunakan sebagai label untuk menjelaskan maksud interaksi pada <i>button</i> . <i>Icon</i> yang ada dapat memberikan kejelasan interaksi pada <i>button</i> .
<i>Button Icon</i>	Digunakan sebagai fungsi dalam interaksi menggunakan <i>button</i> yang terdapat <i>icon</i> . <i>Button icon</i> merupakan <i>button</i> dengan <i>icon</i> sebagai interaksi pada suatu antarmuka.

2. *Perceptual Pattern*

Perancangan *Perceptual Pattern* merupakan perancangan pada elemen yang mendasar sebagai bentuk visual komponen utama. Perancangan *perceptual pattern* terdiri dari warna, spasi dan jenis huruf sebagai penyusun antar muka.

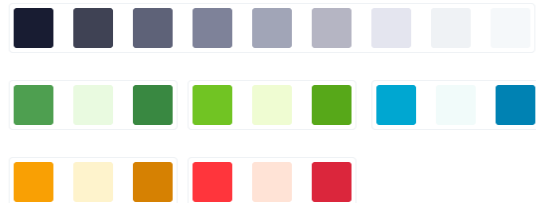
A. Identifikasi Warna

Identifikasi warna terdiri dari kumpulan warna pada Audit antarmuka. Identifikasi warna dapat berupa warna yang dikelompokkan, penentuan pola serta fungsi dari warna, dan pemberian *building block* warna. Gambar 4 merupakan hasil identifikasi warna yang dilakukan pada *website* yang telah ada di CV Digital Akses.



Gambar 4. Identifikasi Warna

Pada proses pengumpulan warna, jika terdapat penggunaan varian warna yang memiliki kesamaan selanjutnya akan dilakukan seleksi warna menjadi warna dasar. Selanjutnya seperti Gambar 5 merupakan hasil identifikasi proses pembentukan warna dilakukan dengan penambahan *shade* dan *tints*. Hasilnya diperoleh level turunan dari warna yang lebih fokus dan tepat.



Gambar 5. Pembentukan Blok Warna

Setelah mengidentifikasi komponen warna, dilakukan perancangan untuk mengelompokkan penggunaan warna. Tabel 5 merupakan analisis pengelompokan *design system* untuk warna yang diusulkan.

Tabel 5. Penggunaan Warna

Kode Hex	Warna	Penggunaan
#4E9F50		<i>Primary</i>
#71C423		<i>Success</i>
#00A7D1		<i>Info</i>
#F9A004		<i>Warning</i>
#FF353C		<i>Danger</i>
#181C32		<i>Neutral</i>

B. Perancangan Jenis Huruf

Pada penelitian ini dalam mengidentifikasi *font* yang akan digunakan pada *design system* menggunakan *font* sistem seperti pada Gambar 6 untuk memberikan pengalaman terbaik di setiap Sistem Operasi. Karena mengikuti standar jenis huruf dari setiap Sistem Operasi masalah kinerja juga berkurang.



Gambar 6. Perancangan Jenis Huruf

3. *Functional Pattern*

Perancangan *functional pattern* dilakukan berdasarkan hirarki pemisahan antar komponen. Pemisahan dilakukan

berdasarkan jenis dan fungsi dari setiap komponen. Tabel 6 merupakan hasil pengelompokan.

Tabel 6. Perancangan *Functional Pattern*

Hirarki	Komponen	Gambar
Form	Button	
	Checkbox	
	Date Picker	
	Number Input	
	Radio	
	Searchbar	
	Select	
	Switch	
	Text Area	
	Form	Text Input
	Time Picker	
	Toogle	
Data Display	Badge	
	Separator	
	Table	
Feedback	Alert	
Navigation	Pagination	
Media	Symbol	
Disclosure	Accordion	
	Tabs	
Overlay	Modal	

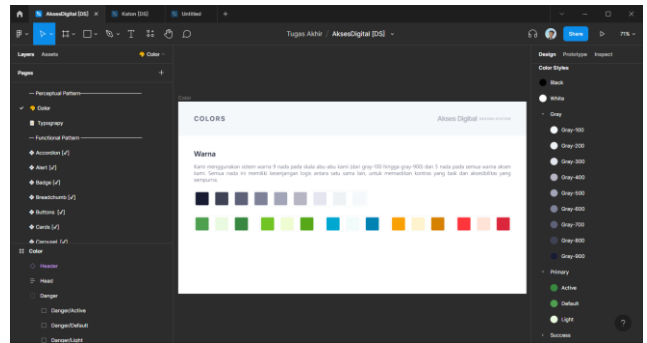
3.6. Implementasi Perancangan

1. Implementasi *Perceptual Pattern*

Implementasi *perceptual pattern* mencakup warna dan font. Adapun penjelasannya adalah sebagai berikut :

A. Warna

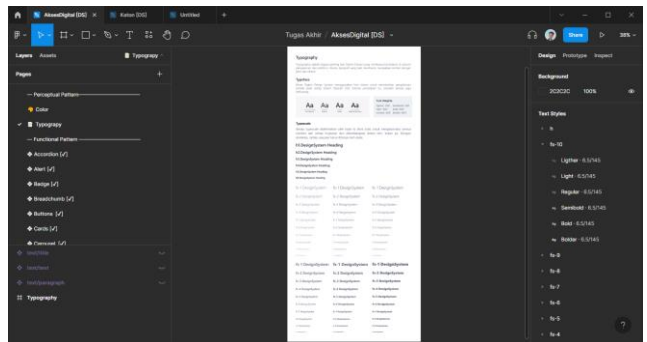
Implementasi sistem warna yang terlihat pada Gambar 7 memiliki kesenjangan logis antara satu sama lain. Hal ini dilakukan untuk memastikan adanya kontras yang baik dan aksesibilitas yang sempurna.



Gambar 7. Implementasi Sistem Warna

B. Typography

Implementasi *typography* seperti pada Gambar 8 dilakukan untuk mendapatkan konsistensi di seluruh pengalaman dan platform. Adanya aturan dalam penentuan tipografi akan memberikan sajian konten dengan jelas dan efisien.



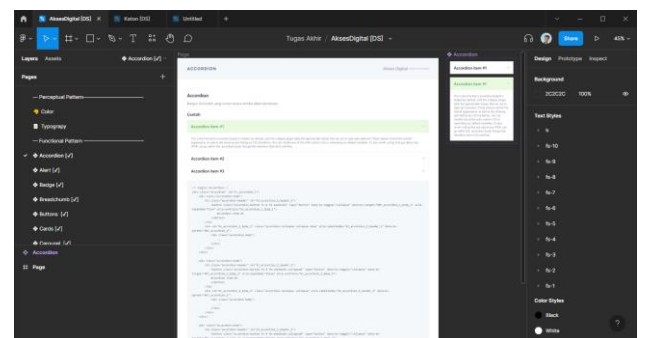
Gambar 8. Implementasi Sistem *Typography*

2. Implementasi *Functional Pattern*

Implementasi dari *functional pattern* dari setiap hasil rancangan yang telah dilakukan terdiri dari :

A. Accordion

Accordion merupakan komponen yang dapat melakukan *collapse* atau memperlihatkan dan menyembunyikan bagian tertentu, seperti terlihat pada Gambar 9.

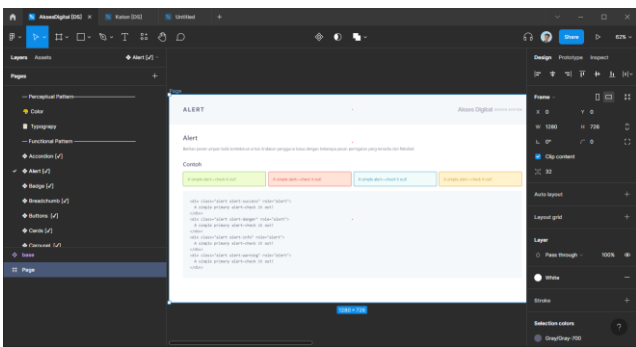


Gambar 9. Implementasi *Accordion* pada FP

B. Alert

Alert diimplementasikan berupa notifikasi pesan umpan balik kontekstual untuk tindakan pengguna biasa dengan

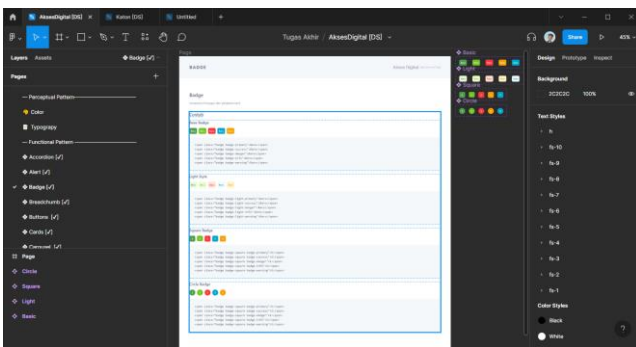
beberapa pesan peringatan yang tersedia dan fleksibel, seperti terlihat pada Gambar 10.



Gambar 10. Implementasi *Alert* pada FP

C. Badge

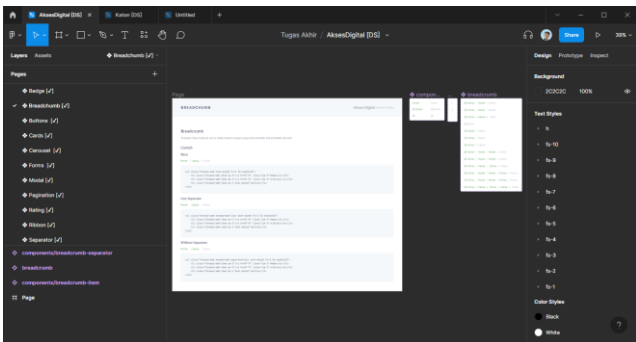
Implementasi *Badge* berupa komponen untuk membuat label, dan perhitungan kecil. Komponen *badge* akan digunakan untuk membuat label, dan menampilkan angka dalam bentuk label. *Badge* diterapkan pada berbagai elemen HTML, seperti dalam *element heading*, tombol, *table* dan lain-lain. Elemen *Badge* seperti terlihat pada Gambar 11.



Gambar 11. Implementasi *Badge* pada FP

D. Breadchumb

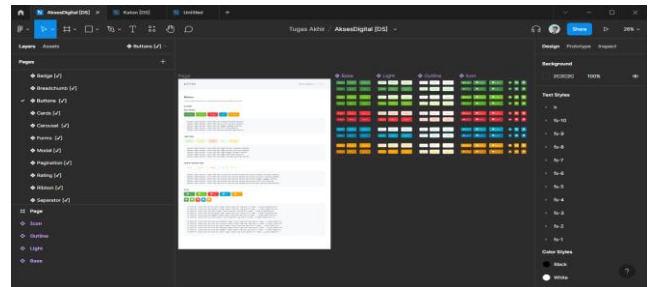
Sebuah komponen penunjuk lokasi halaman saat ini dalam hierarki navigasi yang secara otomatis menambahkan pemisah. Elemen *breadchumb* terlihat pada Gambar 12.



Gambar 12. Implementasi *Breadchumb* pada FP

E. Button

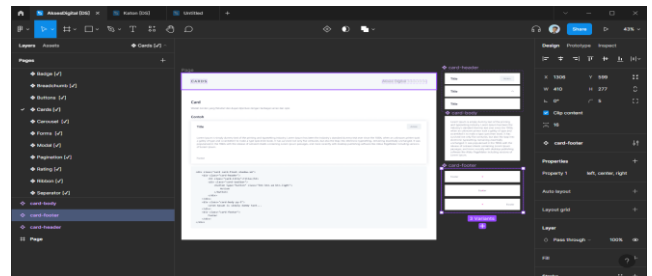
Komponen pada *design system* yang berguna untuk melakukan *action* atau aksi dalam *form* pada sistem. Elemen *button* terlihat seperti pada Gambar 13.



Gambar 13. Implementasi *Button* pada FP

F. Card

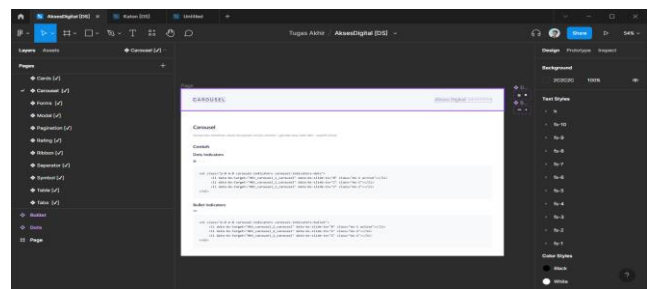
Card merupakan tempat konten yang fleksibel dan dapat diperluas dengan berbagai varian dan opsi. Elemen *Card* terlihat seperti pada Gambar 14.



Gambar 14. Implementasi *Breadchumb* pada FP

G. Carousel

Carousel merupakan hasil implementasi berupa komponen *slideshow*. Elemen *Carousel* terlihat seperti pada Gambar 15.



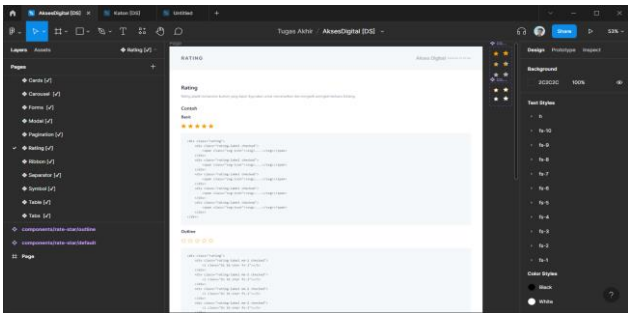
Gambar 15. Implementasi *Carousel* pada FP

H. Form, Modal, serta Pagination

Implementasi *form*, modal, dan *pagination* juga dilakukan dalam implementasi hasil dari perancangan *design system*.

I. Rating

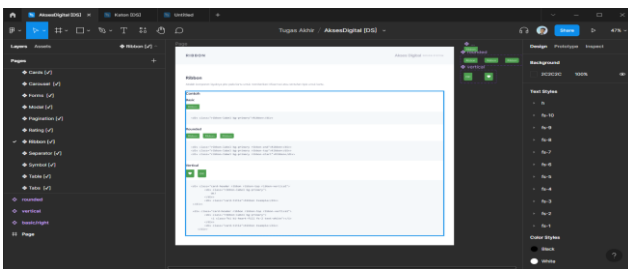
Alert dapat digunakan untuk pesan umpan balik secara kontekstual melalui beberapa pesan peringatan yang tersedia dan *fleksibel*. Implementasi *rating* terlihat seperti pada Gambar 16.



Gambar 16. Implementasi Rating pada FP

J. Ribbon

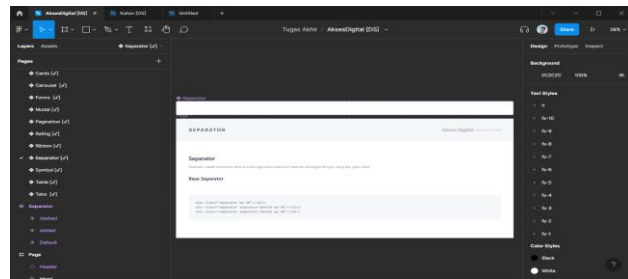
Ribbon merupakan komponen yang dapat digunakan untuk memberikan informasi atau sentuhan tipis untuk kartu. Implementasi Ribbon terlihat seperti pada Gambar 17.



Gambar 17. Implementasi Rating pada FP

K. Separator

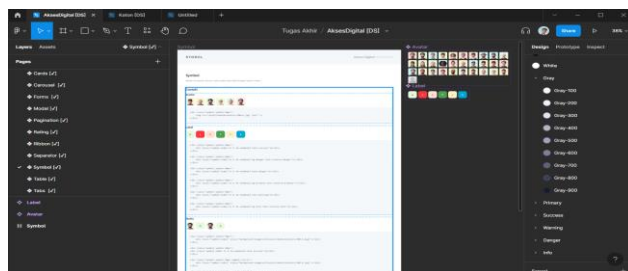
Separator berupa komponen khusus untuk digunakan pada blok halaman perangkat dengan ruang dan garis visual. Implementasi Separator terlihat seperti pada Gambar 18.



Gambar 18. Implementasi Separator pada FP

L. Symbol

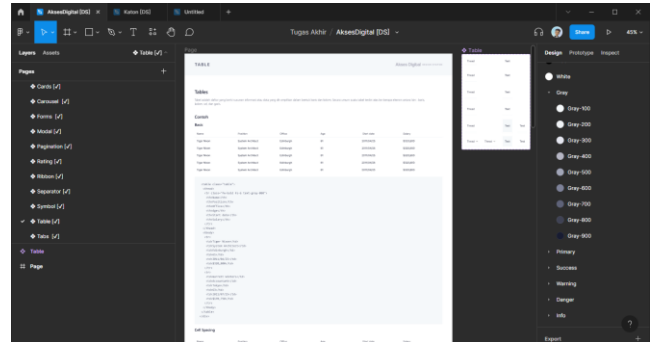
Symbol adalah komponen khusus untuk avatar atau label dengan warna konten. Implementasi Symbol terlihat seperti pada Gambar 19.



Gambar 19. Implementasi Symbol pada FP

M. Table

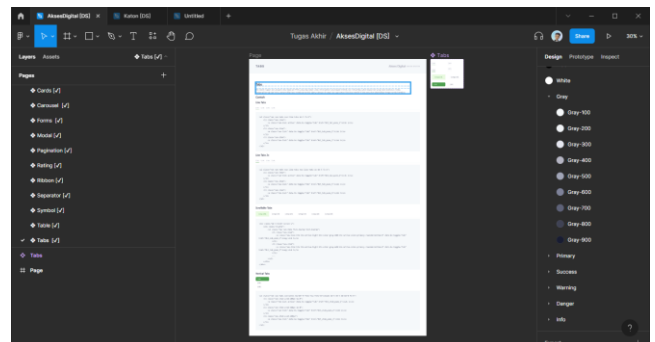
Tabel berupa daftar yang berisi susunan informasi atau data yang ditampilkan dalam bentuk baris dan kolom, Secara umum suatu tabel terdiri atas beberapa elemen antara lain: baris, kolom, sel, dan garis. Implementasi Tabel terlihat seperti pada Gambar 20.



Gambar 20. Implementasi Tabel pada FP

N. Tabs

Tab adalah bagian dari program atau halaman HTML yang digunakan untuk menampilkan satu bagian tertentu dan menyembunyikan bagian lain yang tidak diperlukan. Untuk menampilkan tab dan isinya, pertama kita harus membuat tombol tabnya dengan div atau bisa juga *button*. Kita bisa memilih isi tab yang akan ditampilkan dengan tombol tersebut. Implementasi Tab terlihat seperti pada Gambar 21.



Gambar 21. Implementasi Tab pada FP

3.7. Implementasi Code Library

Setelah implementasi *perceptual pattern* dan *functional pattern* pada *desaign system*. Di-eksekusi ke dalam *code library* menggunakan HTML dan CSS dan disimpan pada <https://aksesdigital-ds.herokuapp.com>

1. Code Library Typography

Hasil dari implementasi *perceptual pattern typography* pada *design system* yang dibuat menjadi *code library* dimana didalamnya terdapat cara untuk menerapkan pada *class html*. Implementasi *Code Library Typography* terlihat seperti pada Gambar 22.

```

$font-sizes: (
  1: $h1-font-size, // 22.75px
  2: $h2-font-size, // 19.50px
  3: $h3-font-size, // 17.55px
  4: $h4-font-size, // 16.25px
  5: $h5-font-size, // 14.95px
  6: $h6-font-size, // 13.95px

```

Gambar 22. Implementasi Code Library Typography

2. Code Library Color

Hasil dari implementasi *perceptual pattern color* pada *design system* yang dibuat menjadi *code library* dimana didalamnya terdapat cara untuk menerapkan warna latar belakang pada class html. Implementasi *Code Library Color* terlihat seperti pada Gambar 23.



Gambar 23. Implementasi Code Library Color

3.8 Pengujian Fungsional

Pengujian fungsional dilakukan dengan metode *test case* dengan menerapkannya menggunakan HTML dan CSS yang sudah dibangun elemen *product-nya*. Metode *test case* merupakan pengujian dengan melakukan verifikasi fitur atau fungsionalitas tertentu dari hasil implementasi aplikasi.

Tabel 7. Pengujian Fungsional Hasil Desain Sistem

Skenario	Data Uji	Hasil Uji	Kesimpulan
Mendisain sebuah halaman <i>auth</i> yang terdiri dari halaman <i>login</i> , <i>register</i> , lupa <i>password</i> , dan <i>password</i> baru.	Halaman login dengan menggunakan <i>pattern library mencakup perceptual pattern</i> dan <i>functional pattern</i> Hifi (<i>High fidelity</i>) halaman <i>register</i> di <i>slicing</i> oleh <i>software engineer</i> kedalam HTML Hifi (<i>High fidelity</i>) halaman lupa <i>password</i> di <i>slicing</i> oleh <i>software engineer</i> kedalam HTML Hifi (<i>High fidelity</i>) halaman <i>password</i> baru di <i>slicing</i> oleh <i>software engineer</i> kedalam HTML	Hasil <i>slicing</i> yang dilakukan <i>software engineer</i> mengikuti hifi halaman <i>password</i> baru yang telah dibuat oleh desainer	100% sesuai dengan hasil rancangan

Dari Tabel 7, hasil pengujian terhadap masing-masing objek hasil perancangan dengan desain sistem menunjukkan hasil 100%. Dalam artian konsep dan penerapan hasil rancangan telah sesuai dengan *framework* yang dibangun dengan desain sistem.

3.9 Pengujian User Acceptance Test.

User Acceptance Test (UAT) merupakan pengujian yang dilakukan secara objektif dengan cara rancangan komponen-komponen untuk aplikasi diuji secara langsung oleh pihak yang menggunakan aplikasi penilaian kinerja ini secara langsung. Bobot nilai pengujian UAT dapat dilihat pada Tabel 8 berikut.

Tabel 8. Bobot Nilai Pengujian UAT

Nilai	Keterangan
4	Sangat Baik
3	Baik
2	Cukup
1	Kurang

Pengujian aplikasi ini dilakukan oleh 26 *user* (termasuk 1 administrator), dengan hasil sebagai berikut:

Tabel 9. Rekapitulasi Penilaian

No	Pertanyaan	Presentase Penilaian (Skor)			
		4	3	2	1
1	Apakah aplikasi mudah digunakan?	80 (77%)	18 (23%)	0 (0%)	0 (0%)
2	Apakah fungsi dan menu pada aplikasi sudah sesuai?	72 (69%)	24 (31%)	0 (0%)	0 (0%)
3	Apakah menu pada aplikasi dapat dimengerti dengan baik?	72 (69%)	24 (31%)	0 (0%)	0 (0%)
4	Apakah fungsionalitas aplikasi sudah menggambarkan kegiatan pelaporan kegiatan yang dilakukan oleh Mahasiswa?	60 (58%)	27 (34%)	4 (8%)	0 (0%)
5	Apakah tampilan antarmuka aplikasi cukup menarik?	72 (69%)	21 (27%)	2 (4%)	0 (0%)

Dari hasil pengujian pada Tabel 9, hasil persentase tingkat kesesuaian sistem yang dihasilkan dari pembagian skor penilaian dengan skor ideal, sebagai berikut :

$$\begin{aligned} \text{Kesesuaian} &= \frac{\text{Total Skor}}{\text{Skor Ideal}} \times 100\% = \\ \frac{((20 \times 4) + (6 \times 3) + \dots + (18 \times 4) + (7 \times 4))}{(26 \times 4) \times 5} &= \frac{476}{520} \times 100 \\ &= 91,54\% \end{aligned}$$

Berdasarkan hasil pengujian menggunakan *User Acceptance Test* (UAT) dengan total 26 responden, diperoleh persentase tingkat kesesuaian sebesar 91,54%. Hasil tersebut menunjukkan bahwa aplikasi yang telah dibangun telah sesuai dengan perancangan dan kebutuhannya.

4. KESIMPULAN

Dari hasil pembahasan dan implementasi *design system* yang telah dilakukan, dapat disimpulkan bahwa elemen *product* yang dibangun *design system* ini adalah *perceptual pattern* dan *functional pattern*. *Perceptual pattern* yang dihasilkan mencakup warna dan *typography*. Sedangkan *functional pattern* yang dihasilkan mencakup *accordion*, *alert*, *badge*, *breadcrumb*, *alert*, *button*, *card*, *carousel*, *form*, *modal*, *pagination*, *rating*, *ribbon*, *separator*, *symbol*, *table*, dan *tabs*. Artefak atau keluaran yang dihasilkan berupa *pattern library*, *style guide*, dan *code library*.

Selain itu untuk pengembangan selanjutnya yaitu: (1) membangun *pattern library* pada platform *mobile* baik Android maupun IOS, (2) membuat *perceptual pattern* yang mencakup panduan pembuatan *icon*, ilustrasi dan *copywriting* pada *design system*, (3) melanjutkan integrasi dari *functional pattern* yang tersedia menjadi sebuah *template* atau halaman.

5. UCAPAN TERIMA KASIH

Terima kasih kepada Allah SWT. Yang telah memberkati dalam penyelesaian jurnal ini, kepada kedua orang tua yang selalu mendukung penulis, kepada kedua dosen pembimbing yang selalu memberikan arahan dengan cermat dan sabar, dan kepada seluruh teman-teman yang selalu memberikan semangat kepada penulis.

DAFTAR PUSTAKA

- [1] A. Kholmatova, "Design Systems. A practical guide to creating design languages for digital products.," *Annals of the CIRP*, vol. 46, no. 1, pp. 75–80, 2017.
- [2] M. Faizal, "Perancangan Simple Design System Canvas sebagai Pedoman Antarmuka Pengguna untuk Startup Indonesia," *Serat Rupa Journal of Design*, vol. 5, no. 1, pp. 108–121, Feb. 2021, doi: 10.28932/SRJD.V5I1.2100.
- [3] I. Hidayanti *et al.*, "Analisis Usability Website Sekolah Menengah Atas," *Jurnal Ilmiah Binary STMIK Bina Nusantara Jaya Lubuklinggau*, vol. 4, no. 2, pp. 86–93, Oct. 2022, doi: 10.52303/JB.V4I2.81.
- [4] D. H. Putra, M. Asfi, and R. Fahrudin, "Perancangan UI/UX Menggunakan Metode Design Thinking Berbasis Web Pada Laportea Company," *Jurnal Ilmiah Teknologi Infomasi Terapan*, vol. 8, no. 1, pp. 111–117, Dec. 2021, doi: 10.33197/JITTER.VOL8.ISS1.2021.730.
- [5] D. D. Aulia, S. Aminah, and D. Sundari, "Perancangan Prototype Tampilan Antarmuka Berbasis Web Mobile Pada Toko Amira Kosmetik," *Jurnal Ilmiah ILKOMINFO - Ilmu Komputer & Informatika*, vol. 5, no. 1, pp. 29–40, Jan. 2022, doi: 10.47324/ILKOMINFO.V5I1.134.
- [6] I. F. Ashari and R. R. Muharram, "Pengembangan Antarmuka Pengguna Kolepa Mobile App Menggunakan Metode Design Thinking Dan System Usability Scale," *JSiI (Jurnal Sistem Informasi)*, vol. 9, no. 2, pp. 168–176, Sep. 2022, doi: 10.30656/JSII.V9I2.4993.
- [7] H. Kurniawan and F. F. Adiwijaya, "Penerapan Desain Sistem Menggunakan Metode Atomic Design Di Universitas Muhammadiyah Sukabumi," *Komputa: Jurnal Ilmiah Komputer dan Informatika*, vol. 10, no. 1, pp. 19–27, Apr. 2021, doi: 10.34010/KOMPUTA.V10I1.6531.
- [8] T. N. Amini, T. N. A. Amini, H. Fabroyir, and R. J. Akbar, "Desain dan Evaluasi Antarmuka Mobile App MyITS Alumni pada Platform Android dan Ios Melalui Pendekatan User-Centered Design," *Jurnal Teknik ITS*, vol. 10, no. 2, pp. A133–A139, Dec. 2021, doi: 10.12962/j23373539.v10i2.63024.
- [9] A. Fauzan, Z. M. Eka, Z. F. Akbar, and K. Fathoni, "Pengembangan Aplikasi Virtual Tour sebagai Media Pengenalan Lingkungan Kampus PENS berbasis Website," *Jurnal Teknologi Terpadu*, vol. 7, no. 1, pp. 23–30, Jul. 2021, doi: 10.54914/JTT.V7I1.341.
- [10] A. P. Wibiwo and F. W. Putra, "Perancangan Visualisasi 'Driver Machine Interface Display' pada Kereta Api berbasis ATO/ATP menggunakan Aplikasi Visual Studio," *Jurnal Teknologi Terpadu*, vol. 7, no. 1, pp. 45–50, Jul. 2021, doi: 10.54914/JTT.V7I1.311.
- [11] C. Boyce and P. Neale, *Conducting in-depth interviews: a guide for designing and conducting in-depth interviews for evaluation input.*, 2nd ed. Watertown: MA: Pathfinder international., 2006.
- [12] B. Frost, *Atomic Design*. Brad Frost, 2016.
- [13] F. R. Kautsar, "Rancang Bangun Sistem Informasi Agroindustri Belimbing Dewa Pemerintah Daerah Kota Depok," S.TP. skripsi, Teknologi Industri Pertanian, IPB (Bogor Agricultural University), Bandung, Indonesia, 2011.



PENGARUH PEMASARAN MEDIA SOSIAL TERHADAP KETERLIBATAN PELANGGAN (SURVEI PADA PENGGUNA HALODOC DI INDONESIA)

Alif Ridha Ramadhani¹, Mochamad Ardan Fauzi², Muhammad Mufti Abdullah³, Syti Sarah Maesaroh⁴, Oding Herdiana⁵

^{1,2,3,4,5} Bisnis Digital, Universitas Pendidikan Indonesia, Kampus Daerah Tasikmalaya
Tasikmalaya, Jawa Barat, Indonesia 46115

alifrmadhani94@upi.edu, moch.ardan.fauzi@upi.edu, muftiabdullah@upi.edu, sytisarah@upi.edu, oding.herdiana@upi.edu

Abstract

This study was conducted to obtain information about the magnitude of the influence of social media marketing on customer engagement. Social media marketing (X) is the independent variable, and customer engagement (Y) is the dependent variable in this study. Followers on the Halodoc Instagram account were selected to be the population in this study, and the sample was taken randomly (simple random sampling), with 120 respondents successfully obtained. The SEM method is used to analyze data with the help of IBM SPSS AMOS 21 software for Windows in the data processing. The results of data processing illustrate that there is an influence of social media marketing on customer involvement of (0.931) with a p-value (0.002) < 0.05. Artificial Intelligence-Analytics for social media is a tool for efforts to increase the competitiveness of official Instagram managers for a business against competitors owned by analyzing content optimization through various services in the form of statistics and metrics from Artificial Intelligence-Analytics providers, which are generally website-based. The author's recommendation for Halodoc is to increase closeness with customers so that their bonds become stronger through planning content that is useful for the community and packaged attractively. Later it is hoped that there will be more user-generated content or posting word-of-mouth recommendation comments from customers to other customers.

Keywords: Customer engagement, E-health, Halodoc, Social Media Marketing, Structural Equation Modeling

Abstrak

Studi ini dilakukan agar memperoleh informasi mengenai besaran pengaruh pemasaran media sosial terhadap keterlibatan pelanggan. Pemasaran media sosial (X) merupakan variabel bebas dan keterlibatan pelanggan (Y) menjadi variabel terikat pada penelitian ini. *Followers* pada akun Instagram Halodoc dipilih untuk menjadi populasi pada penelitian ini dan sampelnya diambil secara acak (*simple random sampling*) dengan berhasil diperoleh sebanyak 120 responden. Metode SEM digunakan untuk menganalisis data dengan bantuan perangkat lunak IBM SPSS AMOS 21 untuk Windows dalam proses olah datanya. Hasil olah data menggambarkan adanya pengaruh pemasaran media sosial terhadap keterlibatan pelanggan sebesar (0,931) dengan nilai $p(0,002) < 0,05$. *Artificial Intelligence-Analytics* untuk media sosial menjadi alat upaya peningkatan daya saing pengelola Instagram resmi untuk suatu bisnis terhadap kompetitor yang dimiliki dengan menganalisis optimasi konten melalui berbagai layanan berbentuk statistik dan metrik dari penyedia. *Artificial Intelligence-Analytics* yang umumnya berbasis *website*. Rekomendasi penulis untuk Halodoc harus meningkatkan keekatan dengan pelanggan supaya bertambah kuat ikatannya melalui perencanaan konten yang bermanfaat bagi masyarakat dikemas secara menarik dan nantinya diharapkan akan bertambah *user generate content* ataupun postingan komentar rekomendasi *word-of-mouth* dari pelanggan ke pelanggan lainnya.

Kata kunci: E-health, Halodoc, Keterlibatan Pelanggan, Pemasaran Media Sosial, Structural Equation Modeling

1. PENDAHULUAN

Social media telah menjadi media komunikasi yang telah mengubah cara kita berkomunikasi dan berinteraksi secara *online*[1]. Laporan statistik menunjukkan bahwa pada Januari 2021, Indonesia memiliki 170 juta pengguna *social media* aktif, peringkat ketiga setelah Cina dan India di

kawasan Asia-Pasifik. Pemasar perlu mengintegrasikan elemen teknologi digital untuk mengalahkan pesaing mereka, karena strategi konvensional saja tidak cukup[2]. *Social media* merupakan alat strategi pemasaran yang digunakan oleh perusahaan. Situs jejaring sosial adalah bagian dari *social media* yang membantu pengguna untuk

saling terhubung dengan pengguna lain. Situs jejaring sosial juga dapat menjadi sarana komunikasi antara perusahaan dan pelanggan, yang dapat menjangkau berbagai pengguna dengan biaya lebih rendah[3]. Menjadi bagian dari kehidupan masyarakat umum yang berguna untuk perusahaan untuk memastikan daya tarik aktivitas pemasaran, reputasi pelanggan dan untuk membangun komunitas[4].

Pemasaran media sosial adalah metode penting untuk bisnis dengan ukuran apa pun untuk mencapai peluang peningkatan penjualan. Pemasaran media sosial terdiri dari aktivitas seperti mengunggah teks, gambar, video dan konten lain yang mendorong keterlibatan publik[5]. *Social media* bukan hanya sebuah platform. Postingan yang diposting oleh pengguna media sosial tentang pengalaman berbelanja dapat berdampak signifikan pada konsumen sehubungan dengan produk dan bisnis[6]. Perusahaan saat ini menyadari manfaat dan keuntungan dari *social media*, yang mengapa perusahaan mencoba hadir melalui pemanfaatan *platform* media sosial contohnya Instagram, Twitter, dan Facebook[7]. *Consumer engagement* adalah sebuah konsep yang dapat dilakukan di berbagai bidang seperti humas, *service* dan pemasaran[8]. *Consumers engage* dengan interaksi merek di berbagai macam platform *social media*. Untuk mengelola keterlibatan ini, merek menggunakan strategi *social media multi platform* dimana merek menciptakan dan menyalurkan konten di berbagai platform. *Consumer engagement* dengan konten yang diunggah di salah satu platform dapat memengaruhi keterlibatan dengan merek di platform lain[9]. Iklan *social media* memperkenalkan pelanggan untuk lebih banyak keterlibatan melalui menyukai, berbagi dan berkomentar[10].

Selama dekade terakhir dalam layanan kesehatan sehari-hari, solusi *information technology* telah melengkapi komunikasi layanan kesehatan tradisional. *Telemedicine* dan informasi kesehatan telah membuat langkah besar menuju peningkatan ketersediaan perawatan dan sumber daya pendidikan untuk penyedia/pasien, meskipun tantangan geografis dan ekonomi yang secara historis menghalangi akses oleh segmen populasi.

Dengan Halodoc, diharapkan kepuasan pelanggan karena kemudahan yang ditawarkan, sehingga pelanggan menjadi pelanggan setia dan loyal. Dan untuk perusahaan itu sendiri, dengan aplikasi Halodoc, identifikasi dan perilaku pelanggan dapat dengan mudah diidentifikasi dan memfasilitasi pengembangan bisnis lebih lanjut. Kepuasan pelanggan merupakan hal yang penting bagi perusahaan jasa karena pelanggan akan menyebarkan kepuasannya kepada pelanggan lain sehingga meningkatkan reputasi merek penyedia jasa. Dalam upaya untuk memuaskan kebutuhan pelanggan, perusahaan dituntut untuk mengetahui kebutuhan maupun keinginan pelanggan yang hampir setiap saat berubah.

Penelitian ini bertujuan untuk menguji variabel yang dimungkinkan berpengaruh pada kepuasan pengguna terhadap aplikasi Halodoc. Serta untuk mengetahui rekomendasi yang dapat meningkatkan penggunaan dan kepuasan pengguna ketika menggunakan aplikasi Halodoc.

2. METODE PENELITIAN

2.1 Metode pengumpulan data, instrumen penelitian, dan metode pengujian

Angket digunakan untuk pengumpulan data penelitian dengan pendekatan kuantitatif yang digunakan pada studi ini. Kuesioner merupakan salah satu metode pengumpulan data yang berisi daftar pertanyaan yang tersusun secara sistematis untuk dijawab oleh responden. Dalam menganalisis data pada penelitian ini menggunakan analisis kuantitatif, analisis kuantitatif adalah proses analisis data yang diperoleh dari hasil responden yang sudah dijawab. Langkah awal dalam analisis data pada penelitian ini adalah melakukan uji validitas instrumen dengan menggunakan *Statistical Package for the Social Sciences (SPSS)* untuk membantu proses analisis data. Selanjutnya melakukan pengukuran korelasi antara jawaban dengan pertanyaan pada uji reliabilitas. Parameter estimasi yaitu *maximum likelihood* melalui serangkaian uji pengukuran model maupun pengujian model *structural* atau hipotesis.

Tabel 1. Hasil Uji Validitas

No	Pertanyaan	<i>r</i> _{hitung}	<i>r</i> _{tabel}	Keterangan
Variety				
1	Tingkat variasi media sosial yang dimiliki <i>e-health platform</i> Halodoc	0,7126	0,1793	Valid
2	Tingkat kesesuaian strategi Pemasaran Media Sosial yang diterapkan <i>e-health platform</i> dalam mengelola akun media sosial resmi Halodoc	0,7048	0,1793	Valid
Intensity				
3	Tingkat efektivitas konten <i>e-health platform</i> di media sosial Halodoc	0,7336	0,1793	Valid
4	Tingkat kemenarikan konten <i>e-health platform</i> untuk pelanggan Halodoc	0,7210	0,1793	Valid
Interaction				
5	Tingkat interaksi pelanggan terhadap <i>e-health platform</i> Halodoc	0,7847	0,1793	Valid
6	Tingkat perhatian pelanggan terhadap <i>e-health platform</i> Halodoc	0,7581	0,1793	Valid
Extraordinary Experience				
7	Tingkat rasa senang pelanggan <i>e-health platform</i> Halodoc	0,7435	0,1793	Valid
8	Tingkat kepuasan dari <i>e-health platform</i> yang digunakan (Halodoc)	0,6737	0,1793	Valid

No	Pertanyaan	rhitung	rtabel	Keterangan
<i>Cognitive Process</i>				
9	Tingkat ingin terus bertahan, terhubung, terus berinteraksi, dan ingin menjadi bagian komunitas dari e-health platform yang digunakan (Halodoc)	0,7307	0,1793	Valid
10	Tingkat koneksi yang bertujuan menjalin hubungan dengan e-health platform yang digunakan (Halodoc)	0,7905	0,1793	Valid

Berdasarkan Tabel 1 skor rhitung setiap butir pertanyaan – pertanyaan dalam pengujian validitas instrumen penelitian hasil yang didapat lebih besar dari 0,1793 maka uji reliabilitas lulus dari perolehan sebanyak 120 responden.

Tabel 2. Hasil Uji Reliabilitas

No	Variabel	Cronbach's Alpha	Keterangan
1	Pemasaran Media Sosial	0,875	Reliabel
2	Keterlibatan Pelanggan	0,947	Reliabel

Berdasarkan Tabel 2 diketahui nilai Cronbach's Alpha dari setiap variabel hasilnya lebih dari 0,6 oleh karena itu pengujian reliabilitas dinyatakan lulus.

Data primer bersumber dari 120 followers Halodoc atau pengguna Halodoc supaya mendapatkan data yang akurat. Data primer yang diperoleh merupakan proses analisis dari kuesioner yang berisi tentang tanggapan konsumen mengenai pemasaran media sosial dan tanggapan mengenai customer engagement. Data sekunder dari literatur ilmiah, penelitian terdahulu yang sesuai sebagai rujukan utama studi ini maupun buku-buku.

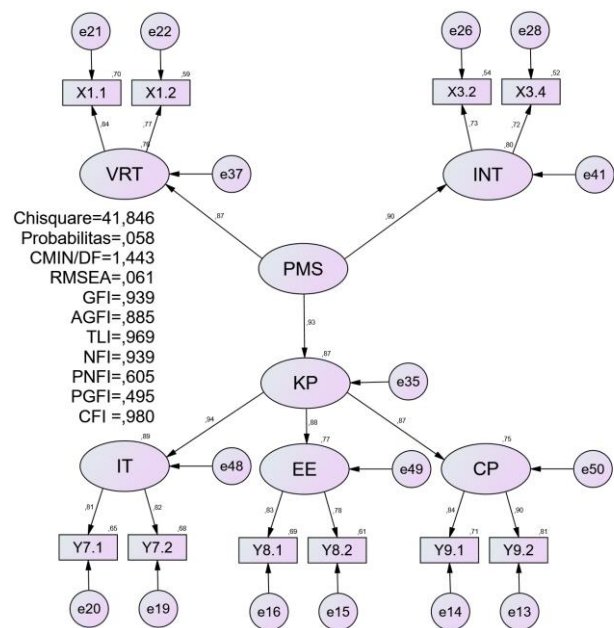
2.2 Tahapan penelitian

Penelitian dilakukan pada awal November menggunakan teknik pengumpulan data menggunakan kuesioner sampai mendapatkan total 120 responden dari berbagai kota di

wilayah Indonesia. Proses analisis data dilakukan pada minggu kedua November di kota Tasikmalaya melalui penyebaran kuesioner secara online. Kuesioner menggunakan semantic differential scale dengan skala 1 sampai 7. Dilanjutkan mengolah data dengan bantuan perangkat lunak IBM SPSS AMOS 21 untuk Windows.

3. HASIL DAN PEMBAHASAN

Gambar 1 memperlihatkan model structural dari kerangka teori yang diusulkan dengan akronim PMS diartikan sebagai variabel laten Pemasaran Media Sosial yang diwakili oleh 2 variabel teramati yaitu akronim VRT sebagai variabel Variety dan akronim INT sebagai variabel Intensity. Selanjutnya akronim KP diartikan sebagai variabel laten Keterlibatan Pelanggan yang diwakili oleh 3 variabel teramati yaitu akronim IT sebagai variabel Interaction, akronim EE sebagai variabel Extraordinary Experience dan akronim CP sebagai variabel Cognitive Process. Model ini memenuhi kriteria yang sudah ditetapkan dan hasilnya sesuai yaitu, nilai p sebesar 0.058, nilai CMIN/DF sebesar 1.443, nilai CFI sebesar 0.980, dan nilai RMSEA sebesar 0.061 .



Gambar 1. Model Structural Uji Kerangka Teori

Tabel 3. Hasil Uji Hipotesis

Hipotesis	Jalur	Path Coef	P	Lower	Upper	Hasil
H ₁	Pemasaran Media Sosial → Keterlibatan Pelanggan	0,931	0,002	0,803	1,04	Diterima
H ₂	Keterlibatan Pelanggan → Extraordinary Experience	0,876	0,002	0,773	0,976	Diterima
H ₃	Keterlibatan Pelanggan → Interaction	0,944	0,002	0,862	1,025	Diterima
H ₄	Pemasaran Media Sosial → Cognitive Process	0,867	0,002	0,752	0,956	Diterima
H ₅	Pemasaran Media Sosial → Intensity	0,897	0,002	0,759	1,037	Diterima
H ₆	Pemasaran Media Sosial → Variety	0,897	0,002	0,76	0,992	Diterima

Berdasarkan Tabel 3 pengujian hipotesis menggunakan metode SEM pada setiap dimensi variabel yang diteliti dapat dilakukan pembahasan yaitu:

- a. Pemasaran media sosial dan keterlibatan pelanggan
Hipotesis pertama meramalkan terdapatnya pengaruh Pemasaran media sosial terhadap keterlibatan pelanggan sebesar (0,931) serta nilai p-nya (0,002) < 0,05 yang membuktikan bahwa hipotesis ini diterima. Dalam selang kepercayaan 95% besar pengaruh Pemasaran media sosial terhadap keterlibatan pelanggan terletak antara 0,803 s.d 1,04. Keberhasilan Pemasaran media sosial terhadap keterlibatan pelanggan menciptakan perilaku non-transaksional yang disebabkan oleh menyebarkan pujian dari *word-of-mouth* melalui postingan komentar media sosial. Hasil ini sejalan dengan penelitian Utami dan Saputri[12].
- b. Keterlibatan pelanggan dan *extraordinary experience*
Hipotesis kedua meramalkan adanya pengaruh keterlibatan pelanggan terhadap *extraordinary experience* sebesar (0,876) serta nilai p (0,002) < 0,05 akhirnya menunjukkan bahwa hipotesis ini diterima. Dalam selang kepercayaan 95% besar pengaruh keterlibatan pelanggan terhadap *extraordinary experience* terletak antara 0,773 s.d 0,976.
- c. Keterlibatan pelanggan dan *interaction*
Hipotesis ketiga meramalkan adanya pengaruh keterlibatan pelanggan terhadap *interaction* sebesar (0,944) serta nilai p (0,002) < 0,05 akhirnya menunjukkan bahwa hipotesis ini diterima. Dalam selang kepercayaan 95% besar pengaruh keterlibatan pelanggan terhadap *interaction* terletak antara 0,862 s.d 1,025.
- d. Keterlibatan pelanggan dan *cognitive process*
Hipotesis keempat meramalkan adanya pengaruh keterlibatan pelanggan terhadap *cognitive process* sebesar (0,867) serta nilai p (0,002) < 0,05 akhirnya menunjukkan bahwa hipotesis ini diterima. Dalam selang kepercayaan 95% besar pengaruh keterlibatan pelanggan terhadap *cognitive process* terletak antara 0,752 s.d 0,956.
- e. Pemasaran media sosial dan *intensity*
Hipotesis kelima meramalkan adanya pengaruh pemasaran media sosial terhadap *intensity* sebesar (0,897) serta nilai p (0,002) < 0,05 akhirnya menunjukkan bahwa hipotesis ini diterima. Dalam selang kepercayaan 95% besar pengaruh pemasaran media sosial terhadap *intensity* terletak antara 0,759 s.d 1,037.
- f. Pemasaran media sosial dan *variety*
Hipotesis keenam meramalkan adanya pengaruh pemasaran media sosial terhadap *variety* sebesar (0,87) serta nilai p (0,002) < 0,05 akhirnya menunjukkan bahwa hipotesis ini diterima. Dalam selang kepercayaan 95% besar pengaruh pemasaran media sosial terhadap *variety* terletak antara 0,76 s.d

0,992. Variasi media sosial yang digunakan oleh Halodoc mampu membantu dalam memperluas jangkauan *customer* sesuai dengan karakteristik masing-masing media sosial yang dikelola[13].

Penggunaan teknologi *Artificial Intelligence-Analitics* dari berbagai ragam penyedia berbasis *website* dapat digunakan untuk mengevaluasi apakah konten Instagram sudah dioptimasi dengan baik sesuai tujuan yang ditetapkan. AI-Analisis ini salah satunya menilai dari sisi statistik kata kunci yang dikemas dalam bentuk tanda pagar apakah sudah tepat atau belum agar jangkauan *audience* dapat dioptimalkan secara lebih luas pada suatu konten yang diposting di luar *followers* yang dimiliki. Teknologi AI-analisis ini dapat juga digunakan sebagai alat untuk menganalisis optimasi Instagram milik kompetitor Halodoc agar setiap pengambilan keputusan Pemasaran Media Sosial dapat lebih baik hasilnya karena salah satu contohnya ialah dalam sisi kata kunci terdapat kata kunci tertentu yang susah untuk ditembus sebab sudah sangat banyak orang/kompetitor yang menggunakan kata kunci tersebut. Selain itu harus disiapkan dana yang memadai karena penyedia AI-Analisis berbasis *website* ini rata-rata mengharuskan membayar sejumlah nominal setelah beralih ke tipe akun premium.

4. KESIMPULAN

Analisis menggambarkan bahwa pemasaran media sosial memiliki pengaruh cukup signifikan terhadap keterlibatan pelanggan maka pada studi ini Halodoc dapat meningkatkan keterlibatan pelanggan melalui pemanfaatan optimasi *sosial media* Instagram agar berkontribusi terciptanya loyalitas pada pelanggan Halodoc.

Studi dilaksanakan pada satu masa yang memungkinkan adanya perubahan kebiasaan pelanggan yang berubah secara cepat dari masa ke masa. Pada penelitian selanjutnya dapat dilakukan dengan memperdalam variabel yang dapat mewakili keseluruhan faktor pemasaran media sosial.

Halodoc harus meningkatkan keeratan dengan pelanggan supaya bertambah kuat ikatannya melalui perencanaan konten yang bermanfaat bagi masyarakat dikemas secara menarik dan nantinya diharapkan akan bertambah *user generate content* ataupun postingan komentar rekomendasi *word-of-mouth* dari pelanggan ke pelanggan lainnya.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada responden penelitian yang telah berpartisipasi dalam pengisian kuesioner.

DAFTAR PUSTAKA

- [1] V. Kumar and P. Nanda, "Social media in higher education: A framework for continuous engagement," *Int. J. Inf. Commun. Technol. Educ.*, vol. 15, no. 1, pp. 109–120, 2019, doi: 10.4018/IJICTE.2019010108.
- [2] G. Cheng *et al.*, "The relationship between csr

- communication on social media, purchase intention, and e-wom in the banking sector of an emerging economy,” *J. Theor. Appl. Electron. Commer. Res.*, vol. 16, no. 4, pp. 1025–1041, 2021, doi: 10.3390/JTAER16040058.
- [3] A. S. Ajina, “The perceived value of social media marketing: An empirical study of online word of mouth in Saudi Arabian context,” *Entrep. Sustain. Issues*, vol. 6, no. 3, pp. 1512–1527, 2019, doi: 10.9770/jesi.2019.6.3(32).
- [4] S. C. Chen and C. P. Lin, “Understanding the effect of social media marketing activities: The mediation of social identification, perceived value, and satisfaction,” *Technol. Forecast. Soc. Change*, vol. 140, no. November 2018, pp. 22–32, 2019, doi: 10.1016/j.techfore.2018.11.025.
- [5] N. Imran and M. Erjumand, “A Study of the Effectiveness of Social Media Marketing Compared to Traditional Marketing Methods,” *RIMS J. Manag.*, vol. 4, no. 2, pp. 1–8, 2020.
- [6] F. J. González-Soriano, P. S. M. Feldman, and J. A. Rodríguez-Camacho, “Effect of social identity on the generation of electronic word-of-mouth (eWOM) on Facebook,” *Cogent Bus. Manag.*, vol. 7, no. 1, 2020, doi: 10.1080/23311975.2020.1738201.
- [7] J. Wang, X. Wang, and Y. Li, “A discrete electronic word-of-mouth propagation model and its application in online social networks,” *Phys. A Stat. Mech. its Appl.*, vol. 527, p. 121172, 2019, doi: 10.1016/j.physa.2019.121172.
- [8] F. Rosado-Pinto and S. M. C. Loureiro, “The growing complexity of customer engagement: a systematic review,” *EuroMed J. Bus.*, vol. 15, no. 2, pp. 167–203, 2020, doi: 10.1108/EMJB-10-2019-0126.
- [9] V. Unnava and A. Aravindakshan, “How does consumer engagement evolve when brands post across multiple social media?,” *J. Acad. Mark. Sci.*, vol. 49, no. 5, pp. 864–881, 2021, doi: 10.1007/s11747-021-00785-z.
- [10] A. A. Alalwan, “Investigating the impact of social media advertising features on customer purchase intention,” *Int. J. Inf. Manage.*, vol. 42, no. April, pp. 65–77, 2018, doi: 10.1016/j.ijinfomgt.2018.06.001.
- [11] M. Deja, D. Rak, and B. Bell, “Digital transformation readiness: perspectives on academia and library outcomes in information literacy,” *J. Acad. Librariansh.*, vol. 47, no. 5, p. 102403, 2021, doi: 10.1016/j.acalib.2021.102403.
- [12] G. R. Utami and M. E. Saputri, “Pengaruh Social Media Marketing Terhadap Customer Engagement dan Loyalitas Merek pada Akun Instagram Tokopedia under a Creative Commons Attribution (CC-BY-NC-SA) 4.0 license CORE View metadata, citation and similar papers at core.ac.uk provided by Electro,” *J. Ris. Manaj. dan Bisnis Fak. Ekon. UNIAT*, vol. 5, no. 2, pp. 185–198, 2020, [Online]. Available: <http://jrmb.ejournal-feuniat.net/index.php/JRMB/article/view/388>.
- [13] L. Lisnawati, R. Hurriyati, and A. W. Al Qorni, “Website Quality and Risk Perception as The Influence of Purchase Intention in E-Commerce Website in Indonesia,” *J. Pendidik. Manaj. Bisnis*, vol. 19, no. 2, pp. 90–99, 2019.

Jurnal Teknologi Terpadu
Volume. 9 No. 1, Juli 2023

Daftar Isi

Kombinasi Linier Target Data untuk Regresi Multitarget menggunakan <i>Principal Component Analysis</i>	01
Yonathan Purbo Santosa	
Sistem Klasifikasi Karakter Kepribadian Siswa Sekolah Dasar berdasarkan Tipologi <i>Hippocrates-Galenus</i> menggunakan Metode Naïve Bayes	10
Muhammad Sabri, Dedy Kasriadi, Irsal, Suci Ramadhani Arifin	
Sistem Informasi Pelayanan Surat Menyurat di Kelurahan Desa Sriwijaya Lampung Tengah	15
Andronias Siregar, Arief Satriansyah, Rachmat Hidayat, Maya Septa Wijaya	
<i>Progressive Web Apps</i>: Pengembangan dan Studi Penerimaan pada Mahasiswa Indonesia menggunakan Scrum dan UTAUT	22
Herman, Frederick	
Perbandingan Model <i>Machine Learning</i> pada Klasifikasi Tumor Otak menggunakan Fitur <i>Discrete Cosine Transform</i>	29
Simeon Yuda Prasetyo, Ghinaa Zain Nabillah	
Analisis <i>Text Mining</i> Klasifikasi Kegiatan Keluarga menggunakan Orange dengan Metode Naïve Bayes	35
Arsya Fathiarahma, Apriade Voutama, Taufik Ridwan, Nono Heryana	
Analisis Sentimen Ulasan Aplikasi MyPertamina pada <i>Google Play Store</i> menggunakan Algoritma NBC	42
Rihan Maulana, Apriade Voutama, Taufik Ridwan	
Pengembangan Aplikasi <i>Emoticon Recognition</i> dan <i>Facial Recognition</i> menggunakan Algoritma <i>Local Binary Pattern Histogram (LBPH)</i> dan <i>Convolutional Neural Network (CNN)</i>	49
Haeruddin, Herman, Patrick Pratama Hendri	
<i>Design System</i> pada Perancangan Antar Muka Perangkat Lunak Sistem Akses Digital	56
Apriansyah Rizqi Setiawan, Marsani Asfi, Agus Sevtiana, Sudadi Pranata, Willy Eka Septian	
Pengaruh Pemasaran Media Sosial terhadap Keterlibatan Pelanggan (Survei pada Pengguna Halodoc di Indonesia)	65
Alif Ridha Ramadhani, Mochamad Ardan Fauzi, Muhammad Mufti Abdullah, Syti Sarah Maesaroh, Oding Herdiana	

Published by :

LPPM STT Terpadu Nurul Fikri
Jln. Raya Lenteng Agung, no. 20, Srengseng Sawah, Jagakarsa,
Jakarta Selatan, DKI Jakarta 12640

Telp. 021 - 786 3191 WhatsApp. 0851 7444 3360

Email : journal@nurulfikri.ac.id

Website : <https://journal.nurulfikri.ac.id/index.php/jtt>

