

ISSN: 2460-8092, E-ISSN: 2548-6551



# SYSTEMIC

INFORMATION SYSTEM AND INFORMATICS JOURNAL

Diterbitkan oleh :  
Program Studi Sistem Informasi  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri (UIN) Sunan Ampel  
Surabaya

**Volume 6 Nomor 1 - Agustus 2020**

# SYSTEMIC

Information System and Informatics Journal

Volume 6 Nomor 1, Agustus 2020

## Pimpinan Redaksi

Indri Sudanawati Rozas (Universitas Islam Negeri Sunan Ampel Surabaya)

## Mitra Bestari

Achmad Solichin (Universitas Budi Luhur)  
Aeri Rachmad (Universitas Trunojoyo Madura)  
Agus Hermanto (Universitas 17 Agustus 1945 Surabaya)  
Ahmad Afif Supianto (Universitas Brawijaya Malang)  
Ahmad Habib (Universitas 17 Agustus 1945 Surabaya)  
Anik Vega Vitianingsih (Universitas Dr. Soetomo Surabaya)  
Devi Karolita (Universitas Palangkaraya)  
Dina Fitria Murad (Universitas Bina Nusantara Jakarta)  
Dwi Puspitasari (Politeknik Negeri Malang)  
Elly Antika (Politeknik Negeri Jember)  
Hanung Prasetyo (Universitas Telkom)  
Hartarto Junaedi (Institut Sains dan Teknologi Terpadu Surabaya)  
Himawan (Sekolah Tinggi Teknologi Informasi NIIT)  
Irwan Alnarus Kautsar (Universitas Muhammadiyah Sidoarjo)  
Moh Noor Al-Azam (Universitas Narotama Surabaya)  
Novi Prastiti (Universitas Trunojoyo)  
Ridha Sefina Samosir (Institut Teknologi dan Bisnis Kalbis)  
Robbi Rahim (Sekolah Tinggi Ilmu Manajemen Sukma)  
Sucipto (Universitas Nusantara PGRI Kediri)  
Sukirman (Universitas Muhammadiyah Surakarta)  
Titin Agustin Nengsih (UIN Sulthan Thaha Saifuddin Jambi)  
Titus Kristanto (Institut Teknologi Telkom Surabaya)  
Uky Yudatama (Universitas Muhammadiyah Magelang)

## Editor

Khalid (Universitas Islam Negeri Sunan Ampel Surabaya)  
Dwi Rolliawati (Universitas Islam Negeri Sunan Ampel Surabaya)  
Noor Wahyudi (Universitas Islam Negeri Sunan Ampel Surabaya)  
M. Andik Izzudin (Universitas Islam Negeri Sunan Ampel Surabaya)  
Mujib Ridwan (Universitas Islam Negeri Sunan Ampel Surabaya)  
Nita Yalina (Universitas Islam Negeri Sunan Ampel Surabaya)  
Ahmad Yusuf (Universitas Islam Negeri Sunan Ampel Surabaya)  
Andhy Permadi (Universitas Islam Negeri Sunan Ampel Surabaya)

# SYSTEMIC

**Information System and Informatics Journal**

**Penerbit**

Program Studi Sistem informasi  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Sunan Ampel Surabaya

**Alamat Redaksi**

Kampus Fakultas Sains dan teknologi  
UIN Sunan Ampel Surabaya  
Jl. Jend. A. Yani 117 Surabaya, 60237  
E-Mail. [systemic@uinsby.ac.id](mailto:systemic@uinsby.ac.id)  
Telp. (031) 8410298, Fax (031) 8413300

---

**SYSTEMIC** merupakan jurnal yang diterbitkan oleh Program Studi Sistem Informasi Fakultas Sains dan Teknologi, UIN Sunan Ampel Surabaya. **SYSTEMIC** diterbit 2 (dua) kali dalam satu tahun pada bulan Agustus dan Desember. Artikel yang dimuat di **SYSTEMIC** berisi pokok bahasan yang terkait dengan aspek pengembangan, kerangka teoritis, implementasi dan pengembangan sistem secara keseluruhan.

---

# SYSTEMIC

## Information System and Informatics Journal

### Daftar Isi

1. Implementasi Kriptografi AES untuk Keamanan Pengiriman Data Internet of Things Menggunakan Web Service Rest pada NodeMCU  
**Afrijal Rizqi Ramadan, Ardito Wahyu Prakoso, Ghifari Dwi C ..... 1 - 6**
2. Perancangan Game Edukasi Kuis Lingkungan Menggunakan Metode Multimedia Development Life Cycle Dan Algoritma Fisher Yates Shuffle Pada Pengacakan Soal  
**Widi Aulia Widi ..... 7 - 12**
3. Rancang Bangun Service Application Program Interface Sistem Machine Learning Klasifikasi Teks Menggunakan Algoritma Support Vector Machine  
**Ottah Hidayatullah, Victor Amrizal, Arini ..... 13 - 21**
4. Analisis Perbandingan Sensitivitas Metode Simple Additive Weighting dan Weighted Product pada Pemilihan Kos di Sekitar UIN Sunan Kalijaga Yogyakarta  
**Irsalina Santi Khasanah, Agus Mulyanto, Muhammad Galih Wonoseto ..... 22 - 28**
5. Sistem Pembelajaran Hukum Baca Al-Qur'an Menggunakan Algoritma LPC dan KNN  
**Hafizh Achmad Dinan, Youllia Indrawaty N, Kurnia Ramadhan Putra ..... 29 - 37**
6. Sistem Penunjang Keputusan Seleksi Atlet Berdasarkan Data Fisik Menggunakan Naïve Bayes  
**yosia halundaka, Din Syamsudin, Aryo Nugroho ..... 38 - 43**
7. Sistem Penunjang Keputusan Pemasaran Produk X Menggunakan Metode K-Means  
**Ach. Syuhbanul Yaumi, Zainul Zulfiqar, Aryo Nugroho ..... 44 - 50**
8. Game Promosi Wisata Kota Malang "Kakang Mbakyu" Dengan Menggunakan Decision Tree dan Hierarchy Finite State  
**Fathurrahman, Yunifa Miftachul Arif ..... 51 - 57**

**SYSTEMIC: Information System and Informatics Journal**

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 1 – Agustus 2020

**Implementasi Kriptografi AES untuk Keamanan Pengiriman Data Internet of Things Menggunakan Web Service Rest pada NodeMCU**Afrijal Rizqi Ramadan<sup>1</sup>, Ardito Wahyu Prakoso<sup>2</sup>, Ghifari Dwi C<sup>3</sup><sup>1,2,3</sup>) UIN Maulana Malik Ibrahim Malang[arizqiramadan@gmail.com](mailto:arizqiramadan@gmail.com)<sup>1</sup>, [arditowahyu19@gmail.com](mailto:arditowahyu19@gmail.com)<sup>2</sup>, [dungeoncity16@gmail.com](mailto:dungeoncity16@gmail.com)<sup>3</sup>**Kata Kunci**IoT,  
REST,  
web service,  
NodeMCU,  
kriptografi,  
AES**Abstrak**

Munculnya istilah Internet of Things pada era ini pada umumnya mempunyai peran positif untuk manusia diberbagai aspek seperti kesehatan, perindustrian, perkotaan dan bahkan pertanian. Namun terdapat masalah yang menjadi ancaman serius dalam lingkungan IoT, yaitu tentang keamanan data. Keamanan data merupakan hal yang diharuskan dalam menyimpan atau mengirim sebuah informasi yang penting. Berbahaya jika data jatuh dan diambil oleh pihak yang nantinya dimanfaatkan dengan tidak bertanggung jawab yang nantinya dapat mengambil bahkan merubah data yang sebenarnya. Oleh karena itu, penggunaan kriptografi AES dibutuhkan untuk menyelesaikan masalah tersebut. AES singkatan dari Advanced Encryption Standard adalah algoritma enkripsi kunci berbentuk simetris yang memiliki proses keamanan data lebih cepat daripada algoritma asimetris. Tujuan penelitian ini yaitu untuk memaparkan cara menerapkan algoritma AES pada perangkat IoT sebagai upaya mengamankan data. Data berupa nilai suhu dan kelembaban yang dienkripsikan dengan algoritma AES sebelum nantinya dikirimkan kepada server basis data dengan wadah berupa jaringan nirkabel atau disebut wifi. Dalam web server data yang terenkripsi dikembalikan dengan bentuk berupa data asli sebelum masuk ke basis data. Hasil dari pengujian ini yaitu mengkonfirmasi bahwa sistem berjalan dengan optimal. Penggunaan algoritma AES dinyatakan berhasil dalam memenuhi tujuan yaitu untuk keamanan serta kerahasiaan data.

**Keywords**IoT,  
REST,  
web service,  
NodeMCU,  
kriptografi,  
AES**Abstract**

The emergence of the term Internet of Things in this era generally has a positive role for humans in various aspects such as health, industry, cities and even agriculture. However, there is a problem that is a serious threat in the IoT environment, namely data security. Data security is something that is required in storing or sending important information. It is dangerous if the data falls and is taken by parties who will be used irresponsibly who can later retrieve and even change the actual data. Therefore, the use of AES cryptography is needed to solve this problem. AES stands for Advanced Encryption Standard is a symmetric key encryption algorithm that has a faster data security process than an asymmetric algorithm. The purpose of this study is to describe how to apply the AES algorithm to IoT devices as an effort to secure data. Data in the form of temperature and humidity values are encrypted by the AES algorithm before being sent to the database server with a container in the form of a wireless network or called wifi. In the web server, encrypted data is returned in the form of the original data before entering the database. The results of this test are to confirm that the system is running optimally. The use of the AES algorithm is declared successful in meeting the objectives, namely for data security and confidentiality.

**1. Pendahuluan**

Internet of Things (IoT) mengalami perkembangan seiring berjalannya perkembangan internet dan mikrokontroler. IoT merupakan konsep dari suatu objek yang bisa melakukan pengiriman suatu data tanpa adanya campur tangan manusia. Namun salah satu unsur penting yang berhubungan dengan pengiriman

data yaitu tentang keamanan. Kerahasiaan dari data merupakan suatu kelengkapan pelayanan agar beberapa atau banyak data yang terkirim tidak sampai dibaca pada pihak ketiga yang tidak berhak menerima. Kriptografi merupakan unsur dari keamanan data yang sering digunakan dan diaplikasikan pada segala bentuk teknologi informasi. Kriptografi mampu merahasiakan segala bentuk data menjadi sebuah data rahasia

(chipertext). Guna menjaga sebuah data yang berada pada perangkat IoT, maka data tersebut harus dilakukan proses enkripsi dan dekripsi yang berfungsi sebagai menjaga kerahasiaan data. Proses enkripsi yaitu suatu proses perubahan dari suatu pesan asli (plain text) menjadi pesan yang rahasia (chipper text), sedangkan proses dekripsi yaitu proses untuk mengembalikan pesan rahasia menjadi pesan dalam bentuk asli. Dalam hal ini, penulis mengaplikasikan kriptografi ke dalam sebuah projek berbasis Internet of Thing (IoT) menggunakan mikrokontroler NodeMCU sebagai perangkat IoT tersebut. Mikrokontroler ini digunakan sebagai media pencari informasi data yang diperoleh dengan bantuan sensor. Lalu data tersebut dikirimkan ke sebuah website pribadi dengan arsitektur RESTful Web Service. Tetapi pada pengiriman data tersebut belum dirahasiakan sehingga data tersebut bisa saja diambil oleh pihak sewenang - wenang. Maka dari itu penulis menggunakan kriptografi yang simetris agar proses keamanan data lebih cepat. Salah satunya yaitu algoritma kriptografi Advanced Encryption Standard (AES).

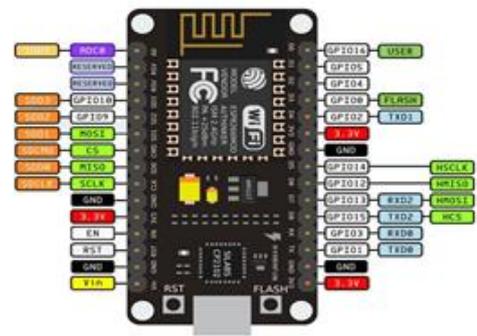
Pada penelitian ini terkonsep pada penerapan algoritma AES guna berfungsi sebagai keamanan sebuah pengiriman data pada perangkat mikrokontroler NodeMCU. Setelah melakukan proses penerapan algoritma AES, selanjutnya adalah dilakukannya pengujian guna menguji waktu proses dan memori yang digunakan pada piranti mikrokontroler yang mendapatkan sebuah atau beberapa kesimpulan berdasarkan pengujian yang dilakukan. Untuk penelitian yang dilakukan oleh penulis diharapkan dapat membuat sebuah langkah awal implementasi algoritma AES dengan proses enkripsi pada pengiriman sebuah data pada lingkungan IoT.

## 2. Teori Dasar

### 2.1 Mikrokontroler NodeMCU V3

NodeMCU V3 adalah sebuah platform Internet of Thing (IoT) yang bersifat opensource, merupakan tahapan perkembangan produk ESP 8266 melalui sebuah firmware menggunakan basis e-Lua. Selain menggunakan bahasa Lua, NodeMCU dapat bekerja sama dengan perangkat lunak bernama Arduino IDE dengan sedikit perubahan di bagian board manager pada Arduino IDE. NodeMCU V3 merupakan perangkat yang kompatibel dengan Bahasa mesin yaitu pemrograman C serta kontak antar perangkat lain atau bisa melalui internet melalui nirkabel. Normalnya NodeMCU mempunyai ukuran panjang 47 mm dan lebar 31 mm. NodeMCU V3 juga termasuk kedalam sebuah mikrokontroler yang mempunyai processor dan memori berukuran kecil sebagaimana spesifikasi dari

NodeMCU V3 dijelaskan dibawah ini:

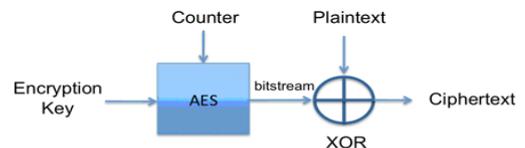


SPEKIFIKASI	NODEMCU V3
Mikrokontroler	ESP8266
Ukuran Board	57 mmx 30 mm
Tegangan Input	3.3 ~ 5V
GPIO	13 PIN
Kanal PWM	10 Kanal
10 bit ADC Pin	1 Pin
Flash Memory	4 MB
Clock Speed	40/26/24 MHz
WiFi	IEEE 802.11 b/g/n
Frekuensi	2.4 GHz - 2.5 GHz
USB Port	Micro USB
Card Reader	Tidak Ada
USB to Serial Converter	CH340G

Gambar 1. Model dan Spesifikasi Mikrokontroler NodeMCU V3

### 2.2 Advanced Encryption Standard (AES)

Advanced Encryption Standard (AES) adalah salah satu algoritma pada kriptografi simetris berfungsi sebagai pengaman pada banyak jenis data. Algoritma ini berupa blok bersifat chipertext simetris yang mempunyai fungsi dapat melakukan kegiatan enkripsi serta kegiatan dekripsi di informasi pada data. Enkripsi dapat merubah sebuah data yang dimana tidak bisa dibaca yang biasa disebut ciphertext; lawan katanya adalah dekripsi yang berguna merubah data ciphertext menjadi bentuk awal data yang dikenal sebagai plaintext. Algoritma AES menggunakan kunci kriptografi blok 128, 192, serta 256 bits guna melakukan proses enkripsi serta proses dekripsi pada sebuah data di blok 128 bits.



Gambar 2. Urutan algoritma AES.

Pada gambar 2 yang tercantum diatas menjelaskan tentang tipe algoritma AES dengan panjang pada kunci, panjang blok serta banyak putaran yang berbeda. Untuk penelitian ini digunakan AES-128 bit dengan jumlah putaran enkripsi sebanyak 10 kali. Terdapat 4 transformasi putaran pada proses enkripsi dan dekripsi [7] :

1. SubBytes

- Berfungsi menukar isi dari komponen data yaitu byte dengan menggunakan tabel substitusi.
- 2. ShiftRows  
Proses pergeseran blok per baris pada state array.
- 3. MixColumn  
Proses mengalikan blok data (pengacakan) di masing-masing state array dengan rumus sebagai berikut:  
 $(x) = \{03\}2 + \{01\}2 + \{01\} + \{02\}$
- 4. AddRoundKey  
Mengombinasikan state array dan round key dengan hubungan XOR.

Pada proses dekripsi algoritma AES :

- a. InvShiftRows  
Melakukan pergeseran bit ke kanan pada setiap blok baris.
- b. InvSubBytes  
Elemen state ditentukan dengan menggunakan sebuah tabel Inverse S-Box.
- c. InvMixColumn  
Kolom pada state dikalikan menggunakan matriks AES.
- d. AddRoundKey  
Mengombinasikan state array dan round key dengan hubungan XOR.

### 2.3 Web Service

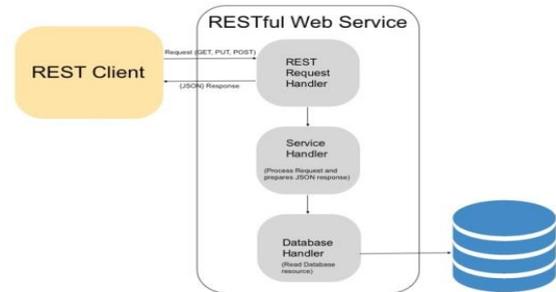
Web Service merupakan sistem software yang dibangun guna mengoptimalkan interaksi dan interoperabilitas antar suatu sistem di suatu jaringan [1]. Web service memiliki interface pada ekstensi yang umumnya dapat dibaca oleh suatu perangkat yang berfungsi sebagai fasilitas untuk menyediakan layanan untuk sistem lainnya agar supaya bisa berinteraksi ke sebuah sistem melewati layanan yang tersedia.

Sistem lain berinteraksi ke web service dengan memakai pesan sejenis SOAP yang pada dasarnya terkirim melewati HTTP dalam bentuk XML [5]. Pada umumnya, web service bukan hanya selalu terdapat pada standar SOAP. Terdapat pustaka yang menjelaskan mengenai web service yang didefinisikan secara umum dengan penjelasan bahwa web service merupakan sebuah aplikasi yang menggunakan jaringan internet untuk mengakses protokol standar internet serta juga untuk menyimpan data informasi berektensi JSON maupun XML, sehingga akses data tersebut dapat digunakan oleh sistem lain walaupun ada perbedaan pada platform, sistem operasi, dan bahasa pemrograman.

### 2.4 Representational State Transfer (REST)

Representational State Transfer atau REST, merupakan salah satu jenis arsitektur untuk

penerapan web service pada konsep perpindahan antar state [2]. State ini menggambarkan jika browser mengakses suatu halaman web atau situs, server akan mentransfer halaman web pada state baru ke browser. Navigasi URL yang telah disiapkan sama halnya seperti mengganti state lama dengan state baru dari halaman situs.



Gambar 3. Kinerja RESTful Web Service  
(Sumber : <https://phppot.com/php/php-restful-web-service/>)

Pada gambar 3, saat REST aktif, dengan bernavigasi dengan pembuatan link-link HTTP disaat proses tertentu, seakan-akan state mengalami perpindahan. Perintah HTTP yang biasa dipakai dalam REST adalah perintah GET, POST, PUT dan DELETE. Lalu balasan dikirim dalam format XML sederhana dan tidak ada protokol pemaketan data, sehingga penerimaan informasi lebih gampang dibaca dan dipasang pada client. Sebutan web service dengan arsitektur REST adalah RESTful web service.

## 3. Perancangan Dan Implementasi Sistem

### 3.1 Perancangan Sistem

Komponen IoT yang digunakan untuk implementasi projek penelitian ini yaitu pada tabel 1 dibawah ini.

Tabel 1:Tabel komponen IoT

Kebutuhan	Keterangan
Sensor DHT11	Modul sensor untuk membaca suhu dan kelembapan
NodeMCU V3 (ESP8266)	Microcontroller yang terintegrasi dengan modul wifi dihubungkan dengan sensor DHT11. Dan NodeMCU juga bertindak sebagai klien REST.
Hosting	Bertindak sebagai server penyimpanan data dan server REST
Wifi Hotspot	Wi-Fi hotspot untuk memberikan koneksi jaringan internet

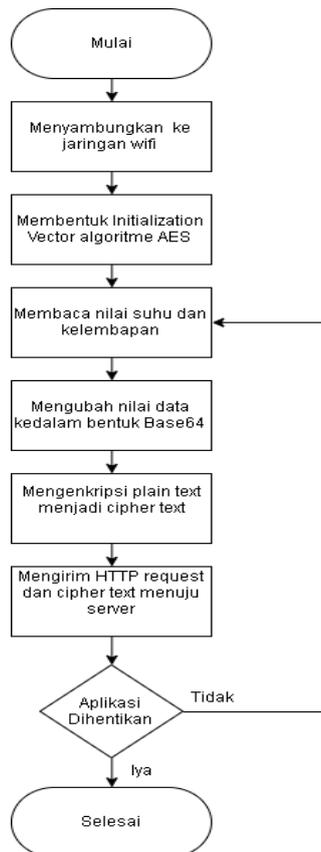
Pada gambar 4, dapat dijelaskan mengenai rancangan konektivitas antar perangkat yang terjadi pada sistem. Pada umumnya sistem tersebut menerapkan arsitektur web service

REST. Dalam pembentukan arsitektur web service yaitu terdapat dua unsur yaitu server dan klien yang saling berkomunikasi melalui jaringan internet. Pada sistem ini, klien akan mengirimkan data dengan request GET kepada server. Lalu server melayani request dan data tersebut. Data kemudian diproses seperti URL yang diakses oleh klien.



Gambar 4. Rancangan proses komunikasi klien server REST.

Gambar 5 menjelaskan tentang tahapan alur proses jalannya data dan kerja sistem pada klien. Sistem dirancang dapat membentuk keystream algoritma AES, tersambung ke router, serta melakukan pembacaan nilai sensor suhu dan kelembapan dari DHT11 lalu data tersebut di enkripsi sebelum dikirim ke server.



Gambar 5. Rancangan alur kerja klien REST.

Gambar 6 merupakan tahapan alur proses kerja sistem di bagian server. Data dari klien terlebih dahulu akan didekripsi. Dilanjutkan nilai

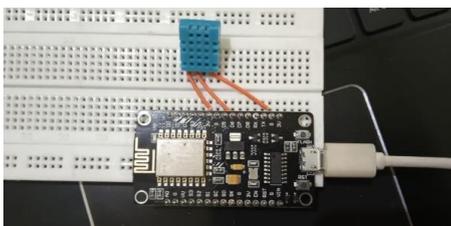
suhu dan kelembapan pada data disimpan oleh basis data MySQL. Terdapat satu tabel yang terdiri dari tiga kolom yaitu kolom id, nama, dan status. Kolom id akan dibentuk secara auto-increment sebagai identitas dari data nama sensor. Kolom nama untuk menyimpan nama sensor yaitu ada sensor suhu dan kelembapan. Kolom status yang bertipe untuk menyimpan data berupa nilai sensor yang diterima oleh server.



Gambar 6. Rancangan alur kerja server REST.

### 3.1 Implementasi

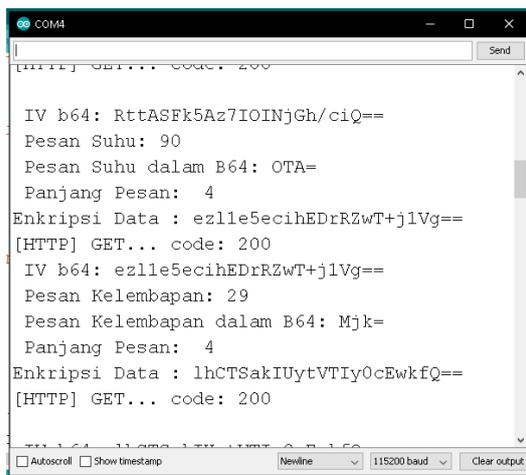
Pada implementasi ini dibutuhkan beberapa komponen diantaranya yaitu sensor DHT11, Projectboard, dan mikrokontroler NodeMCU V3. Sensor DHT11 dan mikrokontroler NodeMCU dihubungkan ke projectboard. Pin kaki positif pada sensor DHT11 dihubungkan dengan kaki 3V pada mikrokontroler NodeMCU V3. Pin kaki negatif pada sensor suhu DHT11 dihubungkan dengan pin GND pada NodeMCU V3. Dan Pin data pada DHT11 dihubungkan dengan pin D6 pada NodeMCU. Kemudian kode program di dalam mikrokontroler diunggah dengan menghubungkan perangkat NodeMCU dengan laptop.



Gambar 7. Implementasi perangkat keras  
Tabel 3. Pengaturan perangkat lunak Arduino IDE.

<b>Board</b>	NodeMCU 1.0 (ESP12-E Module)
<b>Flash size</b>	4M (3M SPIFFS)
<b>Debug port</b>	Serial
<b>Debug level</b>	None
<b>WiFi variant</b>	v2 Prebuilt (MSS=536)
<b>CPU frequency</b>	80 MHz
<b>Upload speed</b>	115200

dan mengunggah kode yaitu menggunakan Arduino IDE. Dalam penggunaannya beberapa konfigurasi juga diperlukan seperti penambahan board pada perangkat lunak Arduino IDE dan merubah pengaturan agar bisa digunakan dan kode program bisa diunggah ke perangkat NodeMCU V3. Berikut adalah tabel pengaturan bisa dilihat pada Tabel 3.

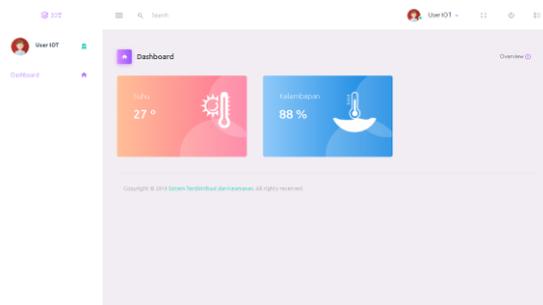


Gambar 8. Tampilan hasil pembacaan serial pada Arduino IDE.

Gambar 8 merupakan output hasil enkripsi dan sistem klien web service REST di dalam Serial Arduino IDE dari gambar tersebut kita dapat menemukan beberapa informasi. Diantaranya yaitu :

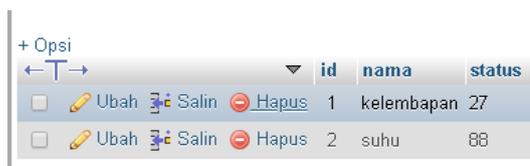
- IV b64 : hasil dari proses pembentukan IV dalam format b64.
- Pesan Suhu: hasil pembacaan nilai sensor suhu
- Pesan Kelembapan : hasil pembacaan nilai sensor kelembapan
- Panjang Pesan : panjang nilai pesan dalam b64, , hasil enkripsi data,

- Enkripsi Data : status request dari klien ke server.



Gambar 9. Tampilan hasil dari implementasi server REST.

Pada Gambar 9 diatas merupakan hasil dari tampilan REST server saat diakses. Kode program web REST server ditulis dengan bahasa pemrograman PHP. Data dalam web REST server didekripsi sampai data berbentuk plaintext dan diubah kembali menjadi bentuk nilai asli yang sebelumnya masih format b64. Dan akhirnya, nilai suhu dan kelembapan dapat disimpan ke dalam tabel basis data MySQLServer seperti pada Gambar 10.



Gambar 10. Tampilan dari hasil penyimpanan data pada MySQL-Server.

## 4. Pengujian

### 4.1 Pengujian Fungsional

Pengujian fungsional berfungsi menguji kesesuaian fungsi-fungsi hasil penerapan rancangan tersebut. Berikut Hasil pengujian fungsional pada Tabel 4.

Tabel 4. Hasil pengujian fungsional sistem.

No	Fungsi	Hasil Uji
1	Menyambungkan ke jaringan wifi	Berhasil
2	Membentuk <i>Initialization Vector</i> algoritma AES	Berhasil
3	Membaca nilai data suhu dan kelembapan	Berhasil
5	Mengubah nilai data kedalam bentuk Base64	Berhasil
6	Mengenkripsi <i>plain text</i> menjadi <i>cipher text</i>	Berhasil
7	Mengirim <i>HTTP request</i> dan <i>cipher text</i> menuju server	Berhasil

8	Menerima HTTP request dan cipher text dari klien	Berhasil
9	Mendekripsi cipher text menjadi plain text	Berhasil
10	Mengubah Base64 menjadi bentuk nilai asli	Berhasil
11	Menyimpan data ke dalam basis data	Berhasil

## 4.2 Pengujian Kinerja

Pengujian kinerja dilakukan sebanyak 50 kali lalu hasil pengujian dirata-rata. Pengujian kinerja ini dilakukan untuk melihat efektifitas memori dan waktu yang dibutuhkan dalam proses enkripsi algoritma AES pada perangkat NodeMCU. Hasil rata-rata penggunaan waktu yang dibutuhkan saat uji enkripsi plain text suhu membutuhkan waktu 5542 mikrodetik sedangkan kelembapan 5868 mikrodetik. Proses enkripsi pada perangkat NodeMCU membutuhkan memori 42984 bita.

## 5. Kesimpulan

1. Hasil pengujian fungsionalitas dan keamanan tersebut dapat disimpulkan bahwa algoritma AES bisa digunakan sebagai keamanan data pada proses transfer data ke server basis data menggunakan NodeMCU.
2. Hasil pengujian kinerja tersebut dapat disimpulkan bahwa algoritma AES selaku kriptografi pada perangkat NodeMCU bersifat valid dan benar.
3. Proses enkripsi data pada perangkat mikrokontroler NodeMCU disinyalir membutuhkan waktu 0,005542 detik, sedangkan memori yang dibutuhkan sebesar 0,8% bagian total memori perangkat. Nilai data pada plain text berbanding lurus dengan nilai waktu dan memori yang dibutuhkan oleh perangkat NodeMCU saat proses enkripsi data.

## Daftar Pustaka

- [1] P.-N. Tan, M. Steinbach, and V. Kumar, "Introduction to Data Mining," 2005.
- [2] J. Han, M. Kamber, and J. Pei, "Data Mining: Concepts and Techniques," *Data Min. Concepts Tech.*, 2012.
- [3] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, 1999.
- [4] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, 2010.

- [5] S. Guha, R. Rastogi, and K. Shim, "Rock: a robust clustering algorithm for categorical attributes," *Inf. Syst.*, vol. 25, no. 5, pp. 345–366, 2000.
- [6] Abdur, R. M., 2016. Middleware for Internet of Things: A Survey. *IEEE Internet of Things Journal*. Vol. 03 No.01.
- [7] Ariyus, D., 2008. Pengantar Ilmu Kriptografi: Teori, Analisis, dan Implementasi. Yogyakarta: ANDI
- [8] Nawir, M., et al., 2018. Internet of Things (IoT): Taxonomy of Security Attacks. Malaysia: University Malaysia Perlis.
- [9] Munir, Rinaldi. 2006. Kriptografi. Penerbit Informatika. Bandung.
- [10] Adhi, J. S. 2005. Kriptografi dengan Algoritma AES untuk Penyandian Data. *Skripsi*. Universitas Kristen Duta Wacana. Yogyakarta.

**SYSTEMIC: Information System and Informatics Journal**

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 1 – Agustus 2020

**Perancangan Game Edukasi Kuis Lingkungan Menggunakan Metode Multimedia Development Life Cycle dan Algoritma Fisher Yates Shuffle Pada Pengacakan Soal**Widi Aulia Rohmah<sup>1</sup><sup>[1]</sup> Universitas Muhammadiyah Sukabumi[widiaten@gmail.com](mailto:widiaten@gmail.com)**Kata Kunci***Game kuis, MDLC, Pengetahuan Alam, Pengkocokan, Fisher Yates Shuffle.***Abstrak**

*Pengetahuan tentang alam atau lingkungan sangat begitu penting dalam kehidupan. Pembelajaran akan ilmu ini pun diperlukan agar siswa mengetahui apapun yang dapat terjadi di dunia ini, terlebih perkembangan teknologi saat ini ilmu dan pengetahuan dapat mudah diberikan melalui berbagai media salah satunya yaitu game yang banyak berkembang dan diminati masyarakat. Dalam pemanfaatannya, game dapat dijadikan sebagai media pembelajaran disekolah sehingga kegiatan belajar dapat menjadi lebih menarik dan menyenangkan. Tujuan lain penelitian ini adalah agar siswa lebih mengetahui dan peduli terhadap alam atau lingkungan. Rancang bangun pembuatan game ini menggunakan metode Multimedia Development Life Cycle (MDLC). Game kuis lingkungan ini merupakan media pembelajaran yang berisi pelajaran ilmu pengetahuan alam tentang sistem organisasi kehidupan, interaksi makhluk hidup dengan lingkungan dan pencemaran lingkungan. Komponen di dalam game ini berupa materi dan permainan kuis serta puzzle. Pertanyaan kuis yang diberikan akan diacak menggunakan algoritma Fisher Yates Shuffle sehingga posisi pertanyaan yang keluar tidak sama.*

**Keywords***Quiz game, Multimedia Development Life Cycle, Natural Knowledge, Shuffle, Fisher Yates Shuffle.***Abstract**

*Knowledge of nature or the environment is very important in life. Learning of this knowledge is also needed so that students know anything that can happen in this world, especially the current technological developments of science and knowledge can be easily provided through various media, one of which is a game that is widely developed and desired by the public. In its use, games can be used as a medium of learning at school so that learning activities can become more interesting and enjoyable. Another goal of this research is for students to know and care more about nature or the environment. The design and construction of this game uses the method of Multimedia Development Life Cycle (MDLC). This environmental quiz game is a learning medium that contains natural science lessons about living organization systems, the interaction of living things with the environment and environmental pollution. Components in this game are in the form of material and quiz and puzzle games. Quiz questions given will be randomized using the Fisher Yates Shuffle algorithm so that the position of the questions that come out is not the same.*

**1. Pendahuluan**

Ilmu pengetahuan alam sangat begitu penting dalam kehidupan sebagai pengetahuan bagaimana kita harus memperlakukan alam. Ilmu ini umumnya membahas tentang semua kejadian yang dapat ditimbulkan dari alam ini termasuk hukum-hukum serta bencana jika kita tidak mampu menjaganya dengan baik. Ilmu yang didapatkan di sekolah ini diharapkan dapat menjadi wadah bagi

siswa untuk mempelajari diri sendiri, makhluk hidup dan alam sekitar. Rasa peduli akan alam sangat begitu penting dan diperlukan, namun rasa itupun seakan mulai hilang dan pudar. Dikutip dari berita Detik.New [1] bahwa konferensi laut PBB tahun 2017 yang diadakan di New York menyebutkan sekitar 1 juta burung laut, 100 ribu mamalia, kura-kura dan ikan mati disebabkan sampah plastik. Disadari ataupun tidak semua

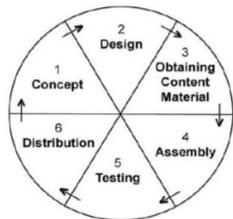
kembali kepada rasa peduli masyarakat akan pentingnya menjaga alam dan isinya. Pengetahuan tentang alam, rasa peduli dan cinta harus ditumbuhkan kembali dalam masyarakat agar alam ini terjaga dan tidak mengalami kerusakan yang parah serta tidak memberikan dampak buruk bagi kehidupan.

Perkembangan teknologi yang semakin pesat menjadikan *game* sebagai salah satu yang paling diminati masyarakat saat ini. Decision Lab And Mobile Marketing Association (MMA) [2] menyebutkan bahwa hasil studi menunjukkan aktifitas yang banyak dilakukan masyarakat adalah bermain *game* yaitu sebanyak 25% dengan durasi rata-rata 53 menit daripada aktifitas lainnya. Dengan peluang itu *game* dapat dijadikan sebagai media pembelajaran yang menarik untuk membantu meningkatkan kembali rasa ingin tahu, motivasi belajar, rasa cinta, serta rasa peduli terhadap alam dan lingkungan. Penanaman karakter ini diperlukan sejak kecil agar ilmu yang diberikan banyak terserap. Dalam prosesnya pembelajaran yang diberikan tidak terlepas dari pertanyaan-pertanyaan agar pengetahuan tersebut bisa menjadi ilmu. Pertanyaan yang diberikan dalam bentuk apapun tentunya harus diatur agar pertanyaan tersebut tidak sama dan membosankan, termasuk dalam *game* kuis lingkungan ini.

Pertanyaan dalam game ini akan diatur dengan mengacak posisi soal yang akan ditampilkan menggunakan algoritma *Fisher Yates Shuffle* agar kuis teracak, tidak sama dan urutan pertanyaan tidak mudah ditebak. Menurut Imam [3] dalam penelitiannya, algoritma Fisher Yates Shuffle dapat menentukan pengacakan yang berubah serta menghasilkan acakan objek yang bermacam-macam.

## 2. Metode Penelitian

Metodelogi yang digunakan yaitu *Multimedia Development Life Cycle* atau yang disebut MDLC dalam pengembangan pembuatan *game* kuis lingkungan ini serta menggunakan algoritma *fisher yates shuffle* untuk pengacakan posisi pertanyaan dalam kuisnya tersendiri.



Gambar 1. Tahapan pengembangan multimedia [4]

### 2.1 Multimedia Development Life Cycle (MDLC)

Menurut Mustika [4], tahapan pengembangan dalam multimedia terdiri dari enam bagian seperti terlihat pada Gambar 1, yaitu pengkonsepan (*concept*), perancangan (*design*), pengumpulan

bahan/materi (*material collecting*), pembuatan (*assembly*), pengujian (*testing*) dan pendistribusian (*distribution*).

#### 2.1.1 Konsep

Tahapan pertama pengkonsepan menentukan tujuan dan manfaat dari *game* kuis lingkungan yaitu sebagai media atau alat bantu dalam pembelajaran ilmu pengetahuan alam materi sistem organisasi kehidupan, interaksi makhluk hidup dengan lingkungan dan pencemaran lingkungan. Target pengguna dalam aplikasi ini adalah siswa-siswi kelas 7 sekolah menengah pertama/ sederajat. *Game* kuis ini terdiri dari tombol materi dan kuis, pengguna dapat memilih tombol materi untuk belajar terlebih dahulu sebelum memulai kuis. Jika pengguna telah siap menjawab kuis maka pengguna dapat memilih langsung tombol kuis tersebut.

Adapun alur dalam permainan kuis yaitu pemain harus menjawab pertanyaan sebanyak 10 pertanyaan dalam 2 menit, setiap pertanyaan memiliki nilai 10 point. Jika pemain dapat menjawab pertanyaan dalam waktu yang telah ditentukan dengan minimal nilai 70 maka pemain dapat melanjutkan permainan selanjutnya dan jika pemain tidak dapat menjawab pertanyaan dalam waktu yang ditentukan dengan nilai seharusnya atau kurang dari 70 maka pemain dapat memilih mengulangi level tersebut atau menutup permainan.

#### 2.1.2 Perancangan

Tahapan berikutnya adalah perancangan, pada tahapan ini membuat perancangan secara rinci tentang *game* kuis lingkungan yang akan dibuat berdasarkan pengkonsepan sebelumnya. Seperti membuat perancangan model aplikasi secara umum, perancangan tampilan yang akan dibuat serta *storyboard* yang menggambarkan alur *game* kuis lingkungan ini.

#### 2.1.3 Pengumpulan bahan

Adapun bahan-bahan untuk membuat *game* kuis lingkungan ini yaitu seperti *background game* berupa alam / lingkungan sekitar, *Icon* untuk menunjukkan tingkatan pada *game*, papan pertanyaan, tombol-tombol, pertanyaan yang akan diberikan dalam *game* kuis lingkungan ini.

#### 2.1.4 Pembuatan

Tahapan pembuatan merupakan tahapan dimana *game* kuis lingkungan dibuat setelah melakukan tahapan-tahapan sebelumnya. Adapun perangkat lunak yang akan digunakan dalam pembuatan aplikasi ini yaitu menggunakan *Adobe Flash* sebagai *software* inti pembuatan *game* kuis.

#### 2.1.5 Pengujian

Pengujian dilakukan setelah tahapan pembuatan selesai dilakukan, dilakukan dengan dua cara yaitu black box testing dan white box

testing. Adapun tujuan dari pengujian ini adalah untuk mengetahui kesalahan yang dapat terjadi pada game kuis lingkungan ini serta memastikan aplikasi yang dibuat telah sesuai dengan apa yang diharapkan.

**2.1.6 Pendistribusian**

Tahapan terakhir yaitu pendistribusian yang merupakan tahapan evaluasi sebuah produk. Selain itu berkas apk dari game yang dibuat disimpan pada media penyimpanan seperti CD, DVD ataupun USB *Flash Drive* kemudian didistribusikan kepada pengguna.

**2.2 Algoritma fisher yates shuffle**

Menurut Hasan [5] dalam penelitiannya, algoritma *Fisher Yates Shuffle* merupakan sebuah algoritma pengacakan yang digunakan untuk menghasilkan permutasi acak dari himpunan terhingga dengan bahasa lain algoritma ini dapat digunakan untuk mengacak himpunan tersebut.

Dalam versi *modern* yang dikenalkan oleh Richard Durstendfeld, algoritma *fisher yates shuffle* bukan lagi dengan cara mencoret angka akan tetapi dengan memindahkan posisi angka yang diambil (k) dengan nilai terakhir (n) yang belum terpilih. Adapun alur algoritma *fisher yates shuffle* sebagai berikut :

1. Tulis angka dari 1 ke N.
2. Ambil angka secara acak, pastikan angka yang belum pernah diambil.
3. Tukar angka (k) dengan nilai terakhir (n) yang belum dipilih.
4. Lakukan kembali langkah 2 dan 3 selama masih ada angka yang belum terambil.
5. Hasil dari pengacakan simpan ditempat lain.

Tabel 1. Contoh pengacakan algoritma fisher yates shuffle [6]

Range	Roll	Scratch	Result
		12345678	
1-8	5	1234867	5
1-7	3	127486	35
1-6	4	12768	435
1-5	5	1276	8435
1-4	2	167	28435
1-3	3	16	728435
1-2	1	6	1728435
<b>Hasil Pengacakan</b>			<b>61728435</b>

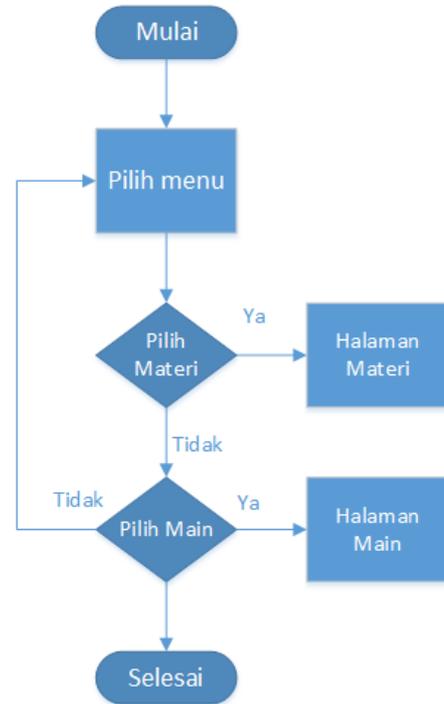
**3. Hasil dan Pembahasan**

Pada hasil dan pembahasan ini berisi *flowchart* skenario umum tentang *game* kuis lingkungan, perancangan dari aplikasi yang akan dibuat dan analisa algoritma *fisher yates shuffle*.

**3.1 Flowchart skenario umum**

Pada *flowchart* ini menggambarkan skenario secara umum dari *game* kuis lingkungan. Aplikasi

ini terdiri dari dua tombol utama, yaitu materi dan kuis. Jika pengguna memilih materi maka aplikasi akan menampilkan halaman menu materi yang dapat dipilih, jika tidak maka pengguna dapat memilih tombol main untuk memulai kuis ataupun keluar dari aplikasi tersebut. Adapun alurnya dapat dilihat pada gambar 2.



Gambar 2. Flowchart skenario umum

**3.2 Tampilan**

Perancangan tampilan yang dibuat diantaranya menu utama, menu materi, menu main, halaman sub materi yang terdiri dari tiga materi yaitu tentang sistem organisasi kehidupan, interaksi makhluk hidup dan lingkungan serta pencemaran lingkungan. Adapun lebih jelasnya sebagai berikut:

**3.2.1 Tampilan menu utama**

Tampilan menu utama dapat dilihat di gambar 3, dimana pada rancangan tampilan menu utama berisi dua tombol yaitu tombol materi dan tombol main. Tombol materi berfungsi untuk menampilkan menu materi. Sedangkan tombol main berfungsi untuk menampilkan menu main.



Gambar 3. Tampilan Menu Utama

### 3.2.2 Tampilan menu materi

Menu materi seperti terlihat di Gambar 4, terdiri dari tiga materi utama yaitu sistem organisasi kehidupan, interaksi makhluk hidup dengan lingkungan dan pencemaran lingkungan.



Gambar 4. Tampilan Menu Materi

### 3.2.3 Tampilan halaman sub materi

Pada tampilan ini terdapat tiga tampilan sub materi tentang pelajaran ilmu pengetahuan alam yaitu :

1. materi sistem organisasi kehidupan, seperti terlihat di Gambar 5.
2. materi interaksi makhluk hidup dengan lingkungan, seperti terlihat di Gambar 6.
3. dan materi pencemaran lingkungan, seperti terlihat di Gambar 7.



Gambar 5. Tampilan Materi Sistem Organisasi Kehidupan



Gambar 6. Tampilan Materi Interaksi Makhluk Hidup



Gambar 7. Tampilan Materi Pencemaran Lingkungan

### 3.2.4 Tampilan menu main

Tampilan menu main, seperti terlihat pada Gambar 8, menampilkan beberapa permainan seperti main kuis sebagai permainan inti dari game kuis lingkungan, main puzzle dan main memasang (drag and drop).



Gambar 8. Tampilan Menu Main

### 3.2.5 Tampilan menu main kuis

Pada tampilan main kuis terdapat tiga kuis sesuai materi yang telah disebutkan dalam sub bab 3.2.3. Tampilan main kuis dapat dilihat gambar 9 berikut.

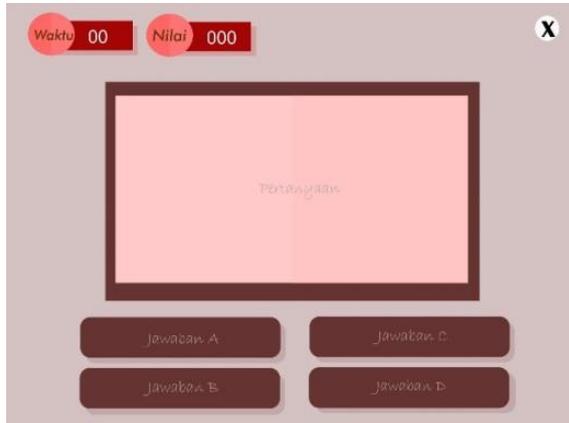


Gambar 9. Tampilan Main Kuis 1

### 3.2.6 Tampilan halaman kuis

Tampilan halaman kuis menggambarkan isi dari main kuis itu sendiri, yaitu desain halaman pertanyaan yang akan diberikan dalam aplikasi (Gambar 10). Berisi papan pertanyaan dan tombol jawaban, nilai yang dapat diperoleh serta waktu yang diberikan untuk dapat menjawab pertanyaan dalam satu level. Adapun alur dalam permainan kuis yaitu pemain harus menjawab pertanyaan sebanyak 10 pertanyaan dalam 2 menit, setiap pertanyaan memiliki nilai 10 point. Jika pemain

dapat menjawab pertanyaan dalam waktu yang telah ditentukan dengan minimal nilai 70 maka pemain dapat melanjutkan permainan selanjutnya (Gambar 10) dan jika pemain tidak dapat menjawab pertanyaan dalam waktu yang ditentukan dengan nilai seharusnya atau kurang dari 70 maka pemain dapat memilih mengulangi level tersebut atau menutup permainan (gambar 11).



Gambar 10. Tampilan Halaman Pertanyaan Kuis



Gambar 10. Tampilan Berhasil

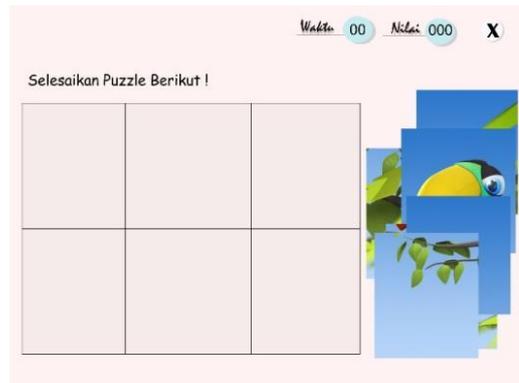


Gambar 12. Tampilan Belum Berhasil

### 3.2.7 Tampilan main puzzle

Selain permainan kuis aplikasi game kuis lingkungan ini terdapat main puzzle. Adapun permainannya seperti pada umumnya, pengguna harus menyusun kembali gambar yang telah

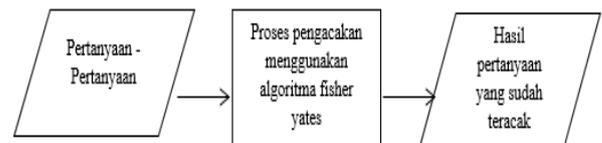
dipotong-potong. Bentuk puzzle dapat dilihat digambar 13.



Gambar 13. Tampilan Main Puzzle

### 3.2.8 Analisa algoritma fisher yates shuffle

Pengacakan yang dilakukan algoritma *fisher yates shuffle* akan selesai jika semua array telah teracak. Adapun gambaran umum langkah-langkah pengacakan pada *game* kuis lingkungan ini menggunakan algoritma *fisher yates shuffle* dapat dilihat pada gambar 14 .



Gambar 14. Gambaran umum pengacakan

## 4. Kesimpulan

Game kuis lingkungan dibuat dengan tujuan dan manfaat sebagai media pembelajaran bagi siswa untuk meningkatkan kembali pengetahuan, rasa ingin tahu, motivasi dalam belajar khususnya ilmu tentang alam. Pengacakan posisi pertanyaan dalam *game* kuis lingkungan menggunakan algoritma *fisher yates shuffle* sehingga pertanyaan yang keluar tidak sama dan teracak.

## Daftar Pustaka

- [1] D. Damarjati, "Data Mengerikan Soal Sampah Plastik di lautan," 24 November, 2018. .
- [2] L. Maulida, "Jumlah Gamer Indonesia," 17 Oktober, 2018. .
- [3] I. Haditama, C. Slamet, and D. Fauzy, "FUZZY TSUKAMOTO DALAM GAME KUIS TEBAK," vol. I, no. 1, pp. 51-58, 2016.
- [4] M. Mustika, E. P. A. Sugara, and M. Pratiwi, "Pengembangan Media Pembelajaran Interaktif dengan Menggunakan Metode Multimedia Development Life Cycle," *J. Online Inform.*, vol. 2, no. 2, p. 121, 2018, doi: 10.15575/join.v2i2.139.

- [5] M. A. Hasan, S. Supriadi, and Z. Zamzami, "Implementasi Algoritma Fisher-Yates Untuk Mengacak Soal Ujian Online Penerimaan Mahasiswa Baru (Studi Kasus : Universitas Lancang Kuning Riau)," *J. Nas. Teknol. dan Sist. Inf.*, vol. 3, no. 2, pp. 291–298, 2017, doi: 10.25077/teknosi.v3i2.2017.291-298.
- [6] D. A. Irawati, A. N. Rahmanto, J. T. Inforasi, and P. N. Malang, "Penerapan Algoritma Fisher Yates Pada Pengacakan Soal Aritmatika," vol. 9, pp. 101–106, 2017.

**SYSTEMIC: Information System and Informatics Journal**

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 1 - Agustus 2020

**Rancang Bangun *Service Application Program Interface* Sistem Machine Learning Klasifikasi Teks Menggunakan Algoritma Support Vector Machine**Ottoh Hidayatullah<sup>1</sup>, Victor Amrizal<sup>2</sup>, Arini<sup>3</sup><sup>1,2,3</sup> Universitas Islam Negeri Syarif Hidayatullah[ottohdayatullah@gmail.com](mailto:ottohdayatullah@gmail.com)<sup>1</sup>, [victor.amrizal@uinjkt.ac.id](mailto:victor.amrizal@uinjkt.ac.id)<sup>2</sup>, [arini@uinjkt.ac.id](mailto:arini@uinjkt.ac.id)<sup>3</sup>**Kata Kunci***Support Vector Machine, Knowledge Acquisition, machine learning***Abstrak**

Data menunjukkan angka yang sangat besar untuk penggunaan Internet di Indonesia. Bidang pendidikan, perpustakaan online adalah upaya memudahkan para peneliti untuk mencari referensi dokumen penelitian. Berdasarkan hasil observasi, UIN Jakarta sudah memiliki repositori dokumen penelitian yang baik, namun pada repositori dokumen penelitian online tersebut belum memenuhi fitur Knowledge Acquisition, kemampuan ini memungkinkan pengguna untuk memperoleh informasi pengetahuan yang tidak mudah langsung di dapat oleh pengguna. Pada penelitian ini dibangun sistem machine learning menggunakan algoritma Support Vector Machine untuk mengkategorikan dokumen berdasarkan bidang penelitian informatika. Penelitian ini juga membangun sistem services API (Application Program Interface) untuk digunakan oleh berbagai macam platform dan lingkungan sistem operasi yang berbeda. Akurasi dari sistem machine learning pada penelitian ini menghasilkan persentase akurasi klasifikasi 73,2% dengan memiliki nilai parameter 0,9. Pada tahap preprocessing pemilihan unigram-bigram adalah yang terbaik, tahap preprocessing yang menggunakan stemming mempengaruhi tingkat klasifikasi sistem machine learning namun mampu meningkatkan hasil akurasi kemampuan. Penggunaan jumlah data mempengaruhi akurasi kemampuan klasifikasi machine learning, terbukti dengan data ditambah menjadi 488 akurasi meningkat menjadi 74,49, dan data bertambah menjadi 492 data maka akurasi meningkat lagi menjadi 77,78%.

**Keywords***Support Vector Machine, Knowledge Acquisition, machine learning***Abstract**

Data shows very large numbers for Internet use in Indonesia. In the field of education, online libraries are an effort to facilitate researchers to search for references to research documents. Based on observations, UIN Jakarta already has a good repository of research documents, but the online research document repository does not fulfill the Knowledge Acquisition feature. This capability allows users to obtain knowledge information that is not easily accessible to users. This research build a machine learning system using the Support Vector Machine algorithm so that the system built can categorize documents based on the informatics research fields. This research also builds a system services API (Application Program Interface) so that data output from machine learning systems can be used by a variety of platforms and different operating system environments. The accuracy of the machine learning system in this study resulted in a percentage of classification accuracy of 73.2% with a parameter value of 0.9. At the preprocessing stage the selection of unigram-bigram is the best in this study. Preprocessing affects the level of classification of machine learning systems. Preprocessing using stemming improves the results of ability accuracy. The amount of data affects the accuracy of the machine learning classification ability, it can be seen when the data is increased to 488 accuracy increases to 74.49. When the experiment was done again so that the data increased to 492 data, the accuracy increased again to 77.78%.

**1. Pendahuluan**

Di bidang pendidikan khususnya di perguruan

tinggi pemanfaatan informasi dapat digunakan sebagai referensi atau penunjang penelitian yang

sedang dilakukan. Hampir seluruh perguruan tinggi memiliki perpustakaan online sebagai *repository* dokumen jurnal dan penelitian ilmiah. Kewajiban untuk memiliki perpustakaan *online* ini di dukung oleh kebijakan pemerintah Indonesia yang tercantum pada Undang - undang Nomor 43 Tahun 2007 Pasal 24 Ayat 3. Berdasarkan observasi pada *website repository Online* pada UIN Jakarta, sistem sudah memiliki kemampuan sebagai *Information Acces* dan *Text Organization*, terdapat fitur *Search Engine* pada *Website repository* perpustakaan *Online* UIN sehingga sistem mampu menyediakan informasi berdasarkan *input* berupa teks dari pengguna. Menurut [1], teks pada pengarsipan dokumen dapat memberikan informasi dengan baik jika memenuhi kemampuan *Information Acces*, *Knowledge Acquisition* dan *Text Organization*. Pengguna perpustakaan dengan mudah menemukan dokumen - dokumen yang tersimpan karena dokumen sudah dikelompokkan dengan susunan struktur yang baik dengan mengelompokkan dokumen tersebut berdasarkan Fakultas dan Jurusan yang ada di UIN Jakarta kemampuan ini sudah memenuhi kriteria *Text Organization*. Untuk kemampuan kriteria sistem manajemen dokumen berupa teks yang harus dimiliki berikutnya adalah kemampuan *Knowledge Acquisition*. Kondisi lain terkait pengguna yang mengakses internet menggunakan berbagai macam *platform* seperti *Web app* dan *Mobile* aplikasi, hal ini menuntut sistem yang mampu mampu terintegrasi dengan *platforms* yang terhubung internet. Untuk memfokuskan penelitian dalam melakukan pengkategorian dokumen ruang lingkup penelitian ini hanya menggunakan dokumen - dokumen jurusan Teknik Informatika sebagai bahan uji coba pembuatan sistem *machine learning* karena bidang tersebut sesuai dengan bidang keilmuan peneliti.

Menurut [2], melakukan klasifikasi data berupa teks dalam bentuk berita menggunakan RapidMiner Studio ke dalam beberapa kategori, TFIDF untuk *feature extraction* dan *support vector machine* sebagai algoritma klasifikasi dengan menggunakan kernel trik RBF dan menghasilkan akurasi yang sangat baik.

Pada [3], melakukan klasifikasi teks Bahasa Indonesia, menggunakan teknik *crawling* dari website berita CNN Indonesia, membandingkan tingkat akurasi algoritma *support vector machine* dan *naïve bayes*, membandingkan pengaruh *feature extraction* SVD (*singular value decomposition*) dan TFIDF (*term frequency inverse document frequency*), hasil penelitiannya bahwa akurasi klasifikasi yang dihasilkan oleh *feature extraction* TFIDF dan algoritma *Mulinomial Naïve bayes* memberikan hasil yang terbaik.

Menurut [4], melakukan klasifikasi menggunakan *support vector machine* kernel *Polynomial* untuk data NUS SMS Corpus dengan jumlah data 40.000, dengan mengambil 6 kelas data, melakukan 2 percobaan yaitu klasifikasi

dengan metode *one-vs-one* yaitu membedakan antara kelas satu dengan semua kelas yang ada yang menghasilkan akurasi 72,6 %, percobaan kedua menambahkan data pada setiap dilakukan proses permodelan, hasilnya adalah tingkat akurasi menghasilkan peningkatan akurasi pada tiap penambahan data.

Sedangkan pada [5], klasifikasi berita bahasa indonesia dengan membandingkan algoritma *support vector machine* dan *K-nearest Neighbo*, menggunakan *stopword* dan TFIDF sebagai *feature extraction*, hasil Algoritma *Support Vector Machine* lebih baik untuk melakukan klasifikasi teks di bandingkan dengan *K-Nearest Neighbor*. Jika dibandingkan antara dua jenis Kernel SVM, maka SVM dengan *Kernel Polynomial* menghasilkan akurasi yang lebih tinggi dengan SVM *Kernel Linier*.

Peneliti [6], klasifikasi berita dari media berita *online* Indonesia bahasa Indonesia, data berasal dari yaitu Kompas.com dengan 12 artikel kategori dengan tiap kategori diambil sebanyak 100 artikel (1200 data), membandingkan algoritma SVM dan NBC untuk mengetahui performa akurasi, *precision*, *recall*, dan *F-Measure*, dan pemrosesan dimana waktu SVM kernel RBF lebih baik dari NBC, peneliti menunjukkan ada 33 berita yang tidak dapat di prediksi dengan baik oleh kedua metode.

Pada [7], fitur *Chi-squared* memberikan hasil terbaik pada algoritma SVM dengan menggunakan metode *kernel*, menggunakan data 457 dokumen tumbuhan obat dan hortikultura yang berasal dari laboratorium temu kembali informasi IPB berbentuk XM. Jumlah data tersebut dibagi menjadi 70% data (320 dokumen) digunakan sebagai data *training* dan 30% (137 dokumen) dijadikan sebagai data uji. Nilai akurasi untuk *Kernel Linier*, *Polinomial*, dan *RBF* adalah 70.8%, 70.8%, dan 73.72%. Dari *feature selection* *Chi-square* menghasilkan akurasi 96.35%, 96.35%, dan 95.62%, sehingga pemilihan fitur *chi-square* membantu klasifikasi SVM dalam mengorganisasikan dokumen secara cepat, efisien, dan dapat meningkatkan kinerja sistem klasifikasi.

Berdasarkan dari studi literatur yang telah digunakan berikut adalah hal-hal yang menjadi fokus pada penelitian kami :

1. Mencari parameter terbaik pada algoritma *support vector machine* sehingga dapat meningkatkan tingkat akurasi klasifikasi
2. Mencari pengaruh proses *preprocessing* *stemming* dan *non-stemming* pada implementasi kinerja *support vector Machine* [8], [9].
3. Mencari pengaruh kinerja algoritma *support vector machine* dalam klasifikasi dengan menggunakan *unigram* dan *bigram* dengan bantuan metode *Chi-square* untuk melihat proses hasil fitur seleksi [10], [11]
4. Mencari pengaruh jumlah data pada *datasets* apakah mempengaruhi kinerja algoritma klasifikasi sistem *machine learning*
5. Merancang *services* API sehingga *machine*

learning dapat terintegrasi dengan platform yang berbeda [12].

6. Sistem *machine learning* ini akan menggunakan metode *support vector machine* dengan kernel trik dan metode *cross validation* sebagai optimasi akurasi.
7. Klasifikasi dokumen pada sistem *machine learning* diterapkan berdasarkan judul dokumen pada jurusan Teknik Informatika.
8. Sistem *machine learning* yang dibangun termasuk ke dalam *supervised learning*.
9. Sistem *Machine Learning* menggunakan data latih dan data uji dari dokumen skripsi repository UIN Jakarta jurusan Teknik Informatika.
10. *Modern Web App* dalam penelitian ini hanya digunakan sebagai bagian uji coba untuk menguji bagaimana *Service* yang dihasilkan dari Sistem *machine learning* dapat digunakan dan berjalan dengan baik [12].
11. Pengkategorian judul skripsi terbagi ke dalam beberapa bidang ilmu informatika yang terdiri dari *Artificial Inteligent, Data Science, Information Sercurity, Internet of Things, Network System, Software Engineer*. Pengkategorian dilakukan secara *manual* data tersebut dikategorikan berdasarkan teori yang penulis ambil dari buku dan jurnal.
12. Tools yang digunakan dalam membangun sistem *Machine Learning* klasifikasi dokumen menggunakan Visual Studio Code sebagai *IDE (Integrated Development Environtment)*
13. Menggunakan bahasa pemrograman Python dengan bantuan ScikitLearn sebagai Library [12], [13], dan [14].
14. Menggunakan MongoDB sebagai *Database* [15].
15. Angular Cli *Framework* untuk membangun *Modern Web App*
16. Menggunakan Docker sebagai teknologi integrasi.

## 2. Metode Penelitian

Dalam melakukan penelitian ini peneliti menggunakan :

1. Metode observasi dan studi pustaka untuk pengumpulan data.
2. Metode pengembangan sistem menggunakan *Continuous ML Model and Control Framework* [16], terdapat 4 fase yaitu *Plan & Acquire, Organize, Analyze Modeling, Deliver*.

### 2.1 Observasi

Peneliti melakukan proses *observasi content analisis* pada repository *online* UIN Jakarta dengan domain <http://repository.uinjkt.ac.id>, untuk identifikasi struktur pola alamat website yang mengacu pada data dokumen skripsi jurusan teknik informatika menggunakan *library python* Scrapy [14].

Table 1. Identifikasi Website Repositori

No	Keterangan	URL
1	URL Halaman utama repository UIN Jakarta	<a href="http://repository.uinjkt.ac.id/dspace/">http://repository.uinjkt.ac.id/dspace/</a>
2	URL repository dokumen skripsi Teknik Informatika	handle/123456789/160
3	URL Parameter untuk mengurutkan data dari yang terbaru	Type=dataaccesssioned &sort_by=2&order=Desc&etal=7&submit browse=Update
4	URL Parameter untuk mendapatkan jumlah data	rpp=480

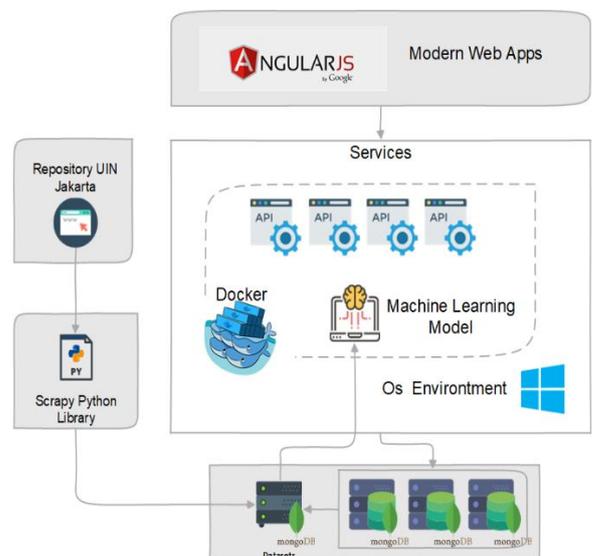
Hasil Identifikasi URL :  
[http://repository.uinjkt.ac.id/dspace/handle/123456789/160/browse?type=dateaccessioned&sort\\_by=2&order=DESC&etal=7&submit browse=Update&rpp=480](http://repository.uinjkt.ac.id/dspace/handle/123456789/160/browse?type=dateaccessioned&sort_by=2&order=DESC&etal=7&submit browse=Update&rpp=480)

## 2.2 Continuous ML Model and Control Framework

### 2.2.1 Plan And Aquire

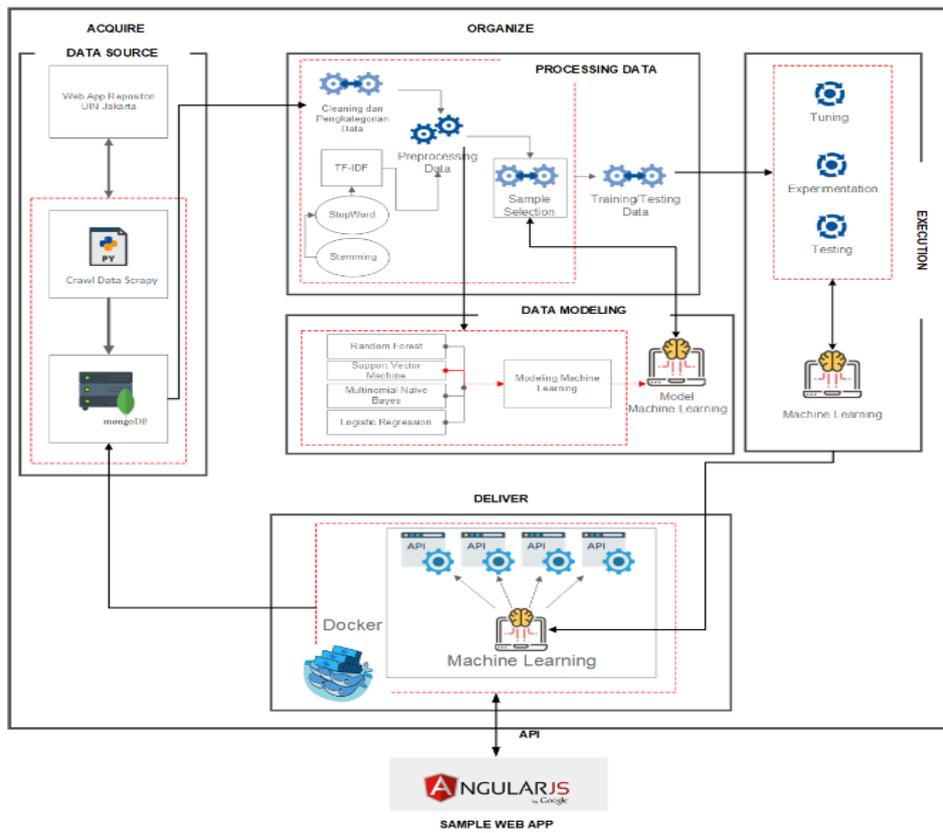
Pada tahap ini mengidentifikasi :

- a. Tujuan dan sumber data.  
 Penelitian ini membangun *sistem Machine Learning* yang mampu melakukan klasifikasi teks dari dokumen skripsi sebagai data. Karena repository UIN Jakarta belum memenuhi *knowledge aquisition*.
- b. Arsitektur sistem dan *platform* pendukung yang digunakan. Gambar 1. adalah arsitektur desain sistem yang dibangun.



Gambar 1. Design Arsitektur Sistem

Perancangan sistem dapat dideskripsikan seperti pada gambar 2.

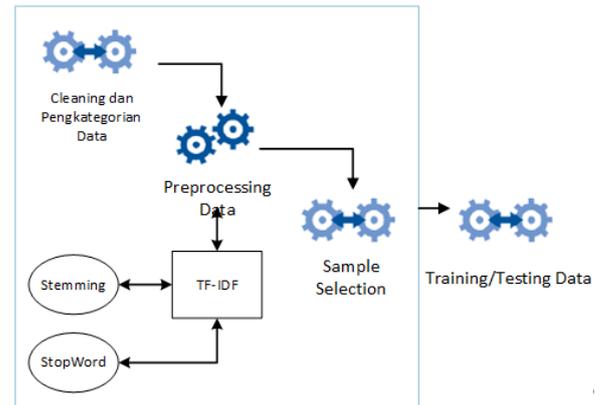


Gambar 2. Arsitektur End to end Machine Learning

2.2.2 Organize

Dari 489 data maka akan dilakukan pengkategorian dan pembersihan data, 80% sebagai data latih dan 20% untuk data uji. Selanjutnya menentukan *Category* sebagai *Class* fitur untuk klasifikasi data secara manual pada dataset. Terdapat 6 *Category* sebagai *Class* fitur klasifikasi data bidang informatika terdiri dari *Software Engineer (SE)*, *Networking System (NS)*, *Internet of Thing (IoT)*, *Artificial Inteligent (Ai)*, *Data Science (DS)* dan *Information Security (IS)*. Untuk mengetahui tingkat akurasi algoritma *Support Vector Machine* dibandingkan dengan algoritma lain seperti *Logistic Regression*, *Multinomial Naive Bayes*, *Random Forest Classifier* dengan bantuan *Library Python SKLearn*. Proses *organize* dibagi menjadi 2 tahap yaitu *feature engineering* dan *model engineering*.

data yang sudah disiapkan pada proses *processing* dan *feature selection* yang bertujuan agar pada tahap selanjutnya sistem dapat fokus untuk mengoptimalkan kinerja sehingga menghasilkan model terbaik.



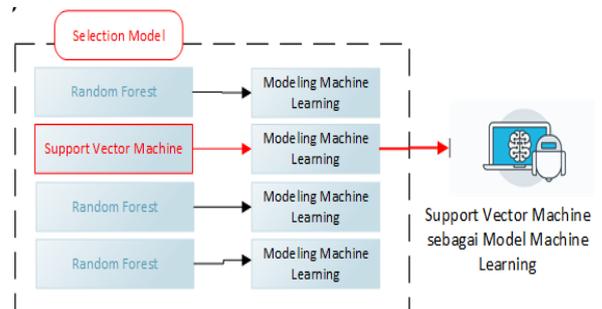
Gambar 3. Proses Tahapan Feature Engineering

2.2.2.1 Feature Engineering (Data Processing)

Tahapan ini melakukan pembersihan data (*cleaning*), pengkategorian data, dan *Preprocessing* yang meliputi *Stemming*, *Stop Words*, *TF-IDF*.

2.2.2.2 Model Engineering (Data Modelling)

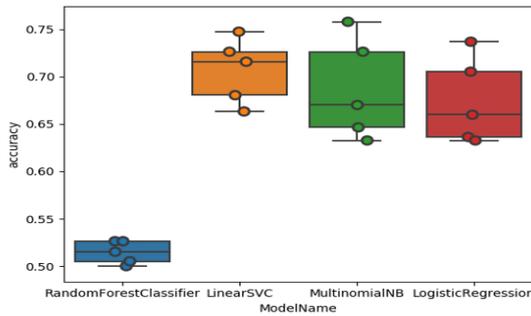
Tahap ini melakukan *model selection*, klasifikasi *Support Vector Machine* dan *Services API Machine Learning*.



Gambar 4. Proses Tahapan Model Engineering

1. Model Selection

Pemilihan algoritma terbaik dilakukan untuk



Gambar 5. Hasil Klasifikasi Dengan Box Whisker Plot

Hasil direpresentasikan dalam diagram *whisker plot*, gambar 5. menunjukkan akurasi 65% hingga 75%, median data berada pada +/- 75% lebih tinggi dibandingkan algoritma lainnya, dan akurasi total dari algoritma SVM mencapai 70%.

### 2. Klasifikasi Support Vector Machine

Klasifikasi menggunakan algoritma *support vector machine multiclass* dan menggunakan *kernel trick* untuk mentransformasikan data ke dalam dimensi *feature space* sehingga mampu mengurangi kesalahan klasifikasi data.

### 3. Services API Machine Learning

Berikut ini adalah API *services machine learning* yang disediakan pada sistem ini :

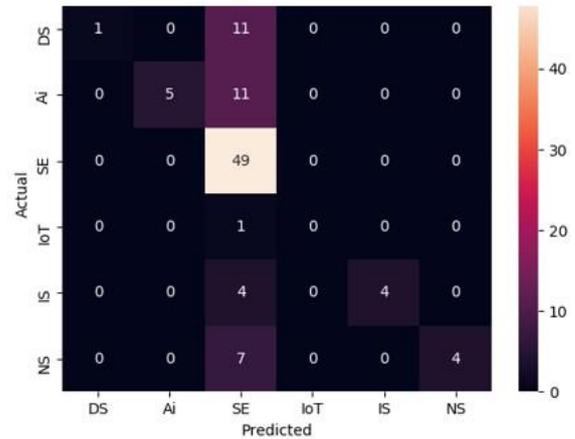
- a) **Insert Data**, bertujuan agar judul skripsi dapat disimpan ke *database* agar jumlah *datasets* pada *database* terus bertambah sehingga *machine learning* dapat meningkatkan proses pembelajaran terhadap data baru. *Service* ini dapat diakses melalui *url /api/InsertData* dengan *method POST*, diuji menggunakan aplikasi *PostMan*, jika sistem mengembalikan nilai *boolean true* maka data berhasil disimpan, jika nilai *false* maka data gagal disimpan di *database*.
- b) **Check Data**, bertujuan mengetahui distribusi data pada *dataset*. *Service* ini dapat diakses melalui *url /api/CheckData*, diuji dengan menggunakan aplikasi *PostMan*. *Method* yang digunakan adalah *GET* sehingga sistem langsung memberikan data dalam bentuk *array* jika ada *request* ke *url /api/CheckData* dengan *method GET*.
- c) **Train Model**, bertujuan agar dapat melakukan pembelajaran pada *datasets* yang terus ditambahkan pada *service InsertData*. Data yang dibutuhkan dilakukan dengan melakukan input nilai parameter C, karena *kernel trik linear* membutuhkan parameter C. *Service* dapat diakses melalui *url /api/train* dengan *method POST*.
- d) **Build Classification**, bertujuan untuk melakukan klasifikasi data yang di *input*. *Service Build Classification* dapat diakses melalui *url /api/BuildClassification* dengan *method POST*.

### 2.2.3 Analyze Modeling

Tahap ini melakukan evaluasi akurasi algoritma *Support Vector*, yaitu dengan melakukan *Experimentation, testing*, dan *tuning*.

#### 1. Model Evaluation

Dari model *machine learning* yang telah dibuat maka dilakukan perhitungan tingkat akurasi. Perhitungan tingkat akurasi ini dilakukan dengan cara model melakukan klasifikasi data uji dengan data latih sebagai referensi. Untuk menghitung tingkat akurasi dengan menggunakan nilai *confusion matrik*, dibuat menggunakan bahasa pemrograman Python.



Gambar 6. Hasil Confusion Matriks

Tabel 2. adalah hasil perhitungan *atribute confusion matrix* :

Table 2. Atribut Confusion Matrik klasifikasi Support Vector Machine

No	Category Data	TP	FP	FN	TN	Jumlah Data
1	DS	1	0	11	85	97
2	Ai	5	0	11	81	
3	SE	49	34	0	14	
4	IoT	0	0	1	96	
5	IS	4	0	4	89	
6	NS	4	0	7	86	

Selanjutnya menghitung nilai *precision, recall, f1-score support* dari setiap kelas.

	precision	recall	f1-score	support
DS	1.00	0.08	0.15	12
Ai	1.00	0.31	0.48	16
SE	0.59	1.00	0.74	49
IoT	0.00	0.00	0.00	1
IS	1.00	0.50	0.67	8
NS	1.00	0.36	0.53	11
avg / total	0.78	0.65	0.59	97
Akurasi Model 64.95 %				

Gambar 7. Hasil Nilai *precision, Recall, F1-score, Support* dan Akurasi

Hasil menunjukkan tingkat akurasi 64,95 %. Sistem mampu mengatasi *error* klasifikasi dilihat dari tinggi nya nilai *precision* pada tiap kelas dan total *precision* cukup baik yaitu 78% namun sistem memiliki tingkat yang sangat buruk pada kelas IoT, karena jumlah data yang sangat sedikit hanya berjumlah 7 data dari total data. Persentase *recall* pada tiap kelasnya belum memiliki nilai yang baik karena hanya kelas SE dan IS yang terbaik di antara lainnya, SE memiliki persentase 100% sedangkan IS memiliki persentase 50%, total persentase keseluruhan *recall* adalah 65%, dapat disimpulkan bahwa kemampuan klasifikasi *machine learning* yang dibangun sudah baik karena sudah melewati angka 50% hanya saja kemampuan klasifikasi *machine learning* masih terbilang kurang baik jika dilihat pada setiap kelasnya.

**2. Performance Engineering**

Tahapan ini untuk melakukan peningkatan kemampuan model *machine learning*, dilakukan dengan proses : *Tuning*, *Experimentation*, dan *Testing*.

**a. Tuning**

Tuning adalah proses pencarian atribut terbaik pada suatu algoritma. Pada percobaan pencarian atribut terbaik, menggunakan perbandingan *kernel trik linear* dan *RBF (radial basis function)*. *Kernel linear* membutuhkan atribut C (konstanta = K), *RBF* membutuhkan atribut C (konstanta = K) dan  $\gamma$  (gamma). Penelitian ini menggunakan *function GridSearchCV* untuk membagi proses kerja kedua kernel, melakukan *cross validation* yaitu membagi *dataset* untuk data latih menjadi beberapa bagian dan pengujian dilakukan berulang kali sebanyak pembagian data latih yang dilakukan.

```
LinearSVC(C=0.9, class_weight=None, dual=True, fit_intercept=True,
intercept_scaling=1, loss='squared_hinge', max_iter=1000,
multi_class='ovr', penalty='l2', random_state=None, tol=0.0001,
verbose=0)
precision recall f1-score support
DS 0.58 0.58 0.54 14
AI 0.70 0.47 0.56 15
SE 0.81 0.88 0.84 52
IoT 1.00 1.00 1.00 2
IS 0.38 0.58 0.43 6
NS 0.75 0.75 0.75 8
avg / total 0.73 0.73 0.73 97
Akurasi Model 73,2 %
----- The SVM model has finished training in 1.0772528375248473 second ----- !
The Script Executed 1.0775812969970783 second !
PS C:\Users\ottho\Learn\ML\Skripsi Machine Learning - Text Classification Repository UIN JKT\ProjectClassific
```

Gambar 8. Hasil Akurasi Model Setelah Tuning

Peneliti mendefinisikan jumlah *cv* adalah 5, sehingga menghasilkan data latih sebanyak 97 data. Untuk *kernel RBF* peneliti mendefinisikan C dan  $\gamma$  (gamma) dari  $1 \times 10^5$  hingga  $1 \times 10^{-5}$ , dan *kernel linear* mendefinisikan C dari  $1 \times 10^3$ , 0,9, 0,7, 0,4, 0,3,  $1 \times 10^{-1}$  hingga  $1 \times 10^{-5}$ . Dari hasil perhitungan kernel linear menghasilkan kernel terbaik untuk digunakan dengan nilai rata-rata 0.705 dan nilai std +/- 0.072 dengan nilai

parameter C = 0.9.

**b. Experimentation**

Peneliti melakukan beberapa hal yang dianggap mempengaruhi tingkat akurasi klasifikasi model *machine learning* dengan memasukkan nilai pembobotan dimana pembobotan tersebut adalah hasil kalkulasi yang dilakukan terhadap hasil *processing* untuk itu peneliti melakukan percobaan pada metode *stemming* pada *preprocessing*. Peneliti melakukan pengujian pembobotan dengan menggunakan *bigram-trigram*, *trigram-quadgram* dan *unigram-bigram* dan bagaimana akurasi model jika data ditambah menjadi 492 data.

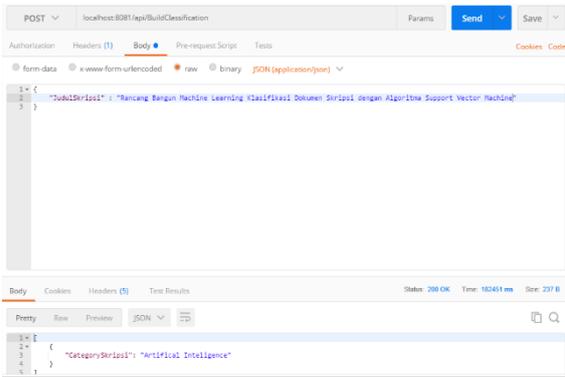
```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL
Jumlah Dataset : 492
Jumlah Data Skripsi Title : 498
Jumlah Data Skripsi Abstrak : 105
Jumlah Data Category :
SE 241
AI 74
NS 63
DS 55
IS 48
IoT 11
Name: SkripsiCategory, dtype: int64
[('SE': 241), ('AI': 74), ('NS': 63), ('DS': 55), ('IS': 48), ('IoT': 11)]
----- Specification Model 1 -----
LinearSVC(C=0.9, class_weight=None, dual=True, fit_intercept=True,
intercept_scaling=1, loss='squared_hinge', max_iter=1000,
multi_class='ovr', penalty='l2', random_state=0, tol=0.0001,
verbose=0)
precision recall f1-score support
DS 0.43 0.67 0.52 9
AI 0.82 0.69 0.75 13
SE 0.85 0.83 0.84 54
IoT 1.00 0.80 0.89 5
IS 0.62 0.83 0.71 6
NS 0.89 0.67 0.76 12
avg / total 0.81 0.78 0.79 99
Akurasi Model 77,78 %
----- The SVM model has finished training in 2.9999999640840187e-07 second ----- !
The Script Executed 59.72084692183577 second !
PS C:\Users\ottho\Learn\ML\Skripsi Machine Learning - Text Classification Repository UIN JKT\ProjectClassific
```

Gambar 9. Hasil Akurasi Model Eksperimen

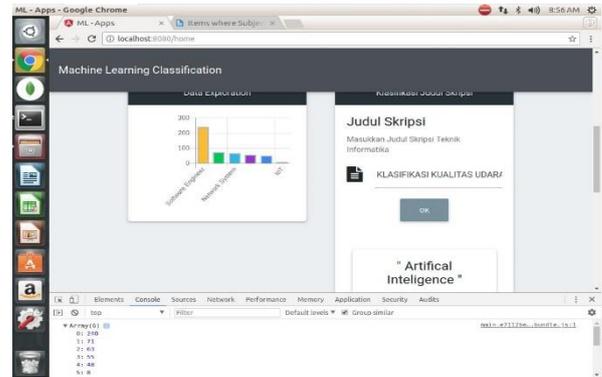
Dari hasil eksperimen diketahui bahwa *preprocessing* dengan menambahkan proses *stemming* sangat berpengaruh, yaitu dari persentase akurasi 73,2% menjadi 77,32% serta nilai *precision*, *recall*, dan *f1-score* memiliki nilai yang seimbang. Hal ini menunjukkan bahwa model *machine learning* mampu menghindari *error* klasifikasi yang baik sebanding dengan kemampuan model *machine learning* untuk melakukan klasifikasi. Akan tetapi proses pembobotan yang melibatkan pemilihan kata menggunakan *bigram-trigram* dan *trigram-quadgram* memberikan dampak pada penurunan akurasi model *machine learning*. Jumlah data mempengaruhi akurasi kemampuan klasifikasi *machine learning*, hal ini ditunjukkan pada pengujian dengan data 488 akurasi meningkat menjadi 74,49 menggunakan kernel *rbf*, nilai C =  $1 \times 10^5$ , dan nilai  $\gamma = 1 \times 10^5$  (gamma). Ketika dilakukan percobaan dengan data bertambah menjadi 492 data maka akurasi meningkat menjadi 77,78% dengan kernel *linear* dan parameter C = 0,9.

**c. Testing**

Tahap ini menguji akses *services API* menggunakan aplikasi *Postman* terhadap aplikasi berbasis web yang telah dibuat pada tahapan *Model Engineering*.



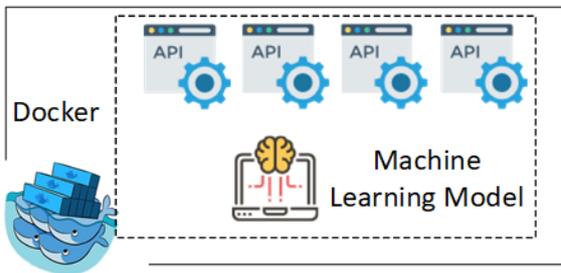
Gambar 10. Testing Output BuildClassification Service API Menggunakan Postman



Gambar 12. Service API CheckData dan BuildClassification Oleh Web App

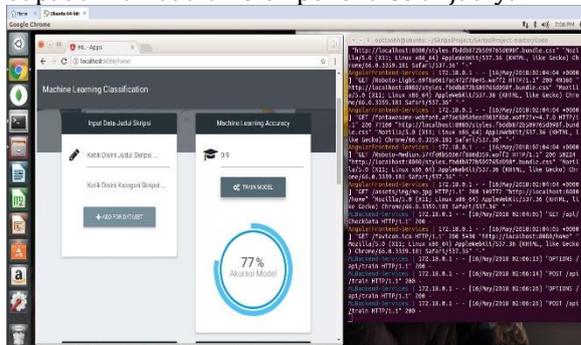
### 2.2.4 Deliver (Deployment)

Tahapan ini bertujuan agar sistem *machine learning* dapat menjadi *services API* (*application program interface*).



Gambar 11. Proses Tahapan Deployment

Dengan menggunakan *Flask Python* untuk membangun *services API machine learning* dan akan menambahkan kemampuan agar dapat berjalan diatas sistem operasi yang berbeda dengan menggunakan *Docker* sehingga sistem *services API* dapat dimanfaatkan oleh pengguna yang memiliki sistem operasi yang berbeda, juga dapat dimanfaatkan oleh peneliti selanjutnya.



Gambar 11. Tampilan Interface Services Sistem Machine Learning dan Modern Web Apps Yang Sudah Terintegrasi.

Pengujian yang dilakukan dengan berbasis web ini yaitu menggunakan aplikasi berbasis web yang memiliki konsep *modern web apps* dengan menggunakan *framework* Angularjs.

### 3. Hasil Dan Pembahasan

Penelitian ini telah berhasil melakukan pengumpulan data, dan melakukan perencanaan (*Plan and Acquire*) dengan melakukan desain arsitektur secara umum dan arsitektur tahapan *End to End ML Architecture*) pada setiap prosesnya, tergambar pada gambar 2.

Selanjutnya pada tahapan *Organize*, telah melakukan persiapan data, data dapat dijadikan obyek klasifikasi pada algoritma SVM, yaitu proses *feature engineering* dan *model engineering*. *Feature engineering* berfungsi untuk melakukan persiapan data mulai dari proses pembersihan data, pengkategorian data, *stemming*, *stopword*, dan pembobotan menggunakan TFIDF (*term frequency inverse document frequency*) sehingga menghasilkan data yang mampu dibaca oleh algoritma yang digunakan. Pada *Model engineering* berfungsi melakukan proses pengujian algoritma *support vector machine* dengan algoritma lain untuk proses klasifikasi yaitu *logistic regression*, *random forest*, *multinomial naive bayes*, yang menghasilkan bahwa algoritma *support vector machine* berada pada akurasi yang baik, terlihat pada gambar 5.

Pada tahapan *Analyze* perancangan telah dapat digunakan untuk mengetahui kemampuan klasifikasi dengan algoritma *support vector machine*. Hal tersebut dilakukan dengan *model evaluation*, telah diketahui seberapa baik sistem *machine learning* menghindari tingkat kesalahan klasifikasi (*precision*), seberapa baik sistem *machine learning* mampu melakukan klasifikasi (*recall*) dan porsi kemampuan keduanya dengan Kemampuan *f1-score*, dengan hasil nilai akurasi *precision* sebesar 78%, *recall* sebesar 65%, *f1-score* sebesar 59%, dan total akurasi sebesar 64,95%. Nilai-nilai tersebut menunjukkan hasil yang baik, terlihat pada gambar 7.

Upaya untuk meningkatkan tingkat akurasi dilakukan pada tahapan *performance engineering* dengan menggunakan 2 proses yaitu *tuning* dan *Experimentation*. Dengan *Tuning* telah dapat melakukan proses pencarian metode dan parameter terbaik yang mampu meningkatkan akurasi model *machine learning*. Pencarian

parameter terbaik menggunakan metode *cross validation* dengan nilai 5, sehingga data latih dibagi menjadi 5 bagian dan setiap bagiannya digunakan oleh model untuk dilakukan proses pengujian terhadap data uji. Hasilnya ditemukan metode terbaik yaitu menggunakan *kernel linear* dengan parameter C yaitu 0,9 dengan metode dan parameter tersebut tingkat akurasi klasifikasi *machine learning* meningkat menjadi 73,2 %. Nilai ini menunjukkan telah ada peningkatan *performance*, terlihat pada gambar 8.

Pada proses *Experimentation* telah mencari kemungkinan yang mampu meningkatkan kemampuan akurasi sistem *machine learning*, yaitu aspek pengaruh proses *stemming* pada *preprocessing* data, dan pengaruh pemilihan n-gram. Pengaruh *stemming* pada persiapan data ditemukan bahwa dengan menggunakan *stemming* akurasi dapat meningkat menjadi *precision* 78%, *recall* 77%, *f1-score* 77%, dan total akurasi menjadi 77,32 % walaupun pemrosesan program berjalan menjadi sangat lambat ketika menggunakan *stemming*, terlihat pada gambar 9.

menggunakan kernel *rbf*, nilai C =  $1 \times 10^5$ , dan nilai  $\gamma = 1 \times 10^5$  (gamma). Ketika dilakukan percobaan kembali sehingga data bertambah menjadi 492 data maka akurasi meningkat lagi menjadi 77,78% dengan kernel *linear* dan parameter C = 0,9, ditunjukkan pada gambar 9.

```
Jumlah Dataset : 488
Jumlah Data Skripsi Title : 486
Jumlah Data Skripsi Abstrak : 185
Jumlah Data Category :
SE 240
AI 73
NS 63
DS 55
IS 48
IoT 9
Name: SkripsiCategory, dtype: int64
[[{"SE": 240}, {"AI": 73}, {"NS": 63}, {"DS": 55}, {"IS": 48}, {"IoT": 9}]
-----Specification Model I-----
SVC(C=100000, cache_size=200, class_weight=None, coef0=0.0,
decision_function_shape='raw', degree=3, gamma=1e-05, kernel='rbf',
max_iter=1, probability=False, random_state=0, shrinking=True,
tol=0.001, verbose=False)
precision recall f1-score support
DS 0.75 0.30 0.43 10
AI 0.59 0.67 0.62 15
SE 0.82 0.91 0.86 45
IoT 1.00 0.67 0.80 3
IS 0.63 0.59 0.62 10
NS 0.63 0.80 0.71 15
avg / total 0.76 0.74 0.73 98

Akurasi Model 74.49 %
----- The SVM model has finished training in 3.99999975698116e-07 second ----- !
The Script Executed 57.68196185119995 second !
PS C:\Users\lototh\Learn\ML\Skripsi Machine Learning - Text Classification Repository UIN JKT\ProjectClassifico
```

Gambar 15. Akurasi Penambahan Data 488

```
'precision', 'predicted', average, warn_for)
precision recall f1-score support
DS 0.00 0.00 0.00 9
AI 0.20 0.18 0.19 11
SE 0.64 0.86 0.73 59
IoT 0.00 0.00 0.00 2
IS 1.00 0.17 0.29 6
NS 0.50 0.20 0.29 10
avg / total 0.52 0.58 0.52 97

Akurasi Model 57.73 %
----- The SVM model has finished training in 7.551626695911917e-07 second ----- !
The Script Executed 0.347456693649292 second !
```

Gambar 13. Akurasi Bigram-Trigram

```
'precision', 'predicted', average, warn_for)
precision recall f1-score support
DS 0.00 0.00 0.00 11
AI 0.00 0.00 0.00 14
SE 0.55 1.00 0.71 51
IoT 0.00 0.00 0.00 2
IS 1.00 0.17 0.29 12
NS 0.50 0.14 0.22 7
avg / total 0.45 0.56 0.42 97

Akurasi Model 55.67 %
----- The SVM model has finished training in 3.775813348649848e-07 second ----- !
The Script Executed 0.9264867305755615 second !
```

Gambar 14. Akurasi Trigram-Quadgram

Pada tahapan *model evaluation* proses *preprocessing* menggunakan *unigram* dan *bigram*, kemudian peneliti melakukan eksperimen terhadap pemilihan n-gram, hasilnya adalah ketika menggunakan *bigram-trigram* nilai total akurasi menurun menjadi 57,73 % dan ketika menggunakan *trigram-quadgram* nilai total akurasi mengalami penurunan lagi sehingga total akurasi menjadi 55,67%, terlihat pada gambar 13 dan gambar 14.

Terkait dengan pengaruh jumlah data yang digunakan, terbukti bahwa jumlah data mempengaruhi nilai akurasi kemampuan klasifikasi *machine learning*, terlihat pada data ditambah menjadi 488 akurasi meningkat menjadi 74,49 %, terlihat pada gambar 15.

Pada tahap *deliver* menunjukkan bahwa sistem telah dapat terintegrasi dengan *platform* serta dapat berjalan pada lingkungan (*environment*) sistem operasi yang berbeda yang menggunakan *library Flask Python*. Data *services API* disajikan dalam format JSON (*javascript object notation*) dengan tujuan agar data dapat diakses melalui HTTP protokol yang sudah tersedia dan dapat diakses oleh berbagai macam jenis *platform*, selain itu juga menggunakan *modern web apps* untuk mengimplementasikan *services API* yang dapat digunakan oleh aplikasi berbasis web. Untuk mendukung sistem yang mampu berjalan di lingkungan (*environment*) sistem operasi yang berbeda menggunakan teknologi Docker, ditunjukkan pada gambar 11 dan gambar 12.

#### 4. Kesimpulan

Penelitian ini berhasil merancang *services API* sistem *machine learning* klasifikasi teks judul skripsi sesuai dengan desain arsitektur sistem yang dirancang pada tahapan pengembangan sistem. Metode dan parameter terbaik adalah algoritma *support vector machine* dengan parameter nilai 0.9, yang menghasilkan persentase akurasi klasifikasi 73,2%.

Pada tahap *preprocessing* pemilihan *unigram-bigram* adalah yang terbaik pada penelitian ini, hal ini dibuktikan ketika menggunakan *bigram-trigram*, dan *trigramquadgram* tingkat akurasi klasifikasi sistem *machine learning* mengalami penurunan. *Preprocessing* data berpengaruh pada tingkat klasifikasi sistem *machine learning*. *Preprocessing* menggunakan *stemming* meningkatkan hasil akurasi kemampuan klasifikasi sistem *machine learning* namun

menghabiskan waktu sekitar 2 menit untuk melakukan pemrosesan program.

Jumlah data juga mempengaruhi akurasi kemampuan klasifikasi *machine learning*, dapat terlihat ketika data ditambah menjadi 488 akurasi meningkat menjadi 74,49, Ketika dilakukan percobaan kembali sehingga data bertambah menjadi 492 data maka akurasi meningkat lagi menjadi 77,78%.

Pada penelitian ini berhasil mengintegrasikan *services API machine learning* pada *platform* yang berbeda dibuktikan dengan melakukan uji coba aplikasi berbasis web dengan memanfaatkan *services API* yang sudah dibangun, juga berhasil menerapkan kemampuan pada sistem *machine learning* klasifikasi dapat berjalan pada lingkungan (*environment*) sistem operasi yang berbeda menggunakan teknologi Docker, dibuktikan dengan sistem dapat berjalan dengan mudah pada sistem operasi Ubuntu Linux.

## Daftar Pustaka

- [1] C. X. Zhai, and S. Massung, "Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining." Penerbit New York : Association for Computing Machinery and Morgan & Claypool , New York. Vol 1, p.7-9, 2016.
- [2] S. M. H. Dadgar, M. S. Araghi, and M. M. Farahani, " A Novel Text Mining Approach Based on TF-IDF and Support Vector Machine for News Classification," IEEE International Conference on Engineering and Technology (ICETECH), 2016, DOI: [10.1109/ICETECH.2016.7569223](https://doi.org/10.1109/ICETECH.2016.7569223)
- [3] R. Wongso, " *News Article Text Classification in Indonesian Language*," Procedia Computer Science 116:137-143 · December 2017, DOI: [10.1016/j.procs.2017.10.039](https://doi.org/10.1016/j.procs.2017.10.039)
- [4] M. Kretchmar, and Y. Zhao, "Text Message Authorship Classification Using kernel Support Vector Machines," IEEE International Conference on Computational Science and Computational Intelligence, 2014, DOI: [10.1109/CSCI.2014.121](https://doi.org/10.1109/CSCI.2014.121)
- [5] S. N. Asiyah, and K. Fithriasari, "Klasifikasi Berita online menggunakan Support Vector Machine dan K-nearest Neighbor," Jurnal Sains dan Seni ITS, Vol 5, No 2 (2016) DOI: [10.12962/j23373520.v5i2.16643](https://doi.org/10.12962/j23373520.v5i2.16643), [http://ejournal.its.ac.id/index.php/sains\\_seni/article/view/16643](http://ejournal.its.ac.id/index.php/sains_seni/article/view/16643)
- [6] D. Ariadi, and K. Fithriasari, "Klasifikasi Berita Indonesia menggunakan Naive Bayesian dan Support Vector Machine dengan Confix Stripping Stemmer," Jurnal Sains dan Seni ITS, Vol 4, No 2, 2015, DOI: [10.12962/j23373520.v4i2.10966](https://doi.org/10.12962/j23373520.v4i2.10966)
- [7] A. D. Putri, and J. Adisantoso, " Klasifikasi dokumen teks menggunakan metode Support Vector Machine dengan pemilihan fitur CHI-Square," 2013, <http://repository.ipb.ac.id/handle/123456789/65199>
- [8] A.Z. Arifin, I.P.A.K. Mahendra, and H.T. Ciptaningtyas, "Enhanced Confix Stripping Stemmer and Ants Algorithm for Classifying News Document in Indonesian Language," Proceeding of International Conference on Information & Communication Technology and Systems (ICTS), p. 149-157, 2009
- [9] A. D. Tahitoe, and D. Purwitasari, "Implementasi Modifikasi Enhanced Confix Stripping Stemmer Untuk Bahasa Indonesia dengan Metode Corpus Based Stemming," Institut Teknologi Sepuluh Nopember (ITS) – Surabaya, 60111, Indonesia, 2010, <http://digilib.its.ac.id/public/ITS-Undergraduate-14255-paperpdf.pdf>
- [10] D.N. Chandra, G. Indrawan, & I.N. Sukajaya, "Klasifikasi Berita Lokal Radar Malang Menggunakan Metode Naive Bayes Dengan Fitur N-Gram," Jurnal Ilmiah Teknologi Dan Informasi ASIA (JITIKA), 10(1),p.11-19. 2016 <http://lp3m.asia.ac.id/wpcontent/uploads/2016/02/2.-jurnal-Denny.pdf>
- [11] E. Indrayuni, M. Wahyudi, S. Informasi, J. Selatan, I. Komputer, & J. Selatan, "Penerapan Character N-Gram Untuk Sentiment Review Hotel Menggunakan Algoritma Naive Bayes," *Konfrensi Nasional Ilmu Pengetahuan dan Teknologi (KNIT)* (pp. 88-93). 2015.
- [12] G. Dwyer, S. Aggarwal, J. Stouffer, "Flask: Building Python Web Services. Packt, ", Publishing Ltd, Birmingham, 2017
- [13] Haroon, Danish and Karachi, "Python Machine Learning Case Studies, " Penerbit Apress, Pakistan, 2017
- [14] R. Lawson, " Web Scraping with Python. Packt," Publishing Ltd, Birmingham, 2015
- [15] MongoDB Compass. 2016. Dari : <https://www.mongodb.com/blog/post/getting-started-with-mongodb-compass>. Diakses pada 25 Mei 2018 pukul 20.30 WIB.
- [16] Gartner, "Preparing and Architecting for machine learning," 2017. [https://www.gartner.com/binaries/content/assets/events/keywords/catalyst/catus8/preparing\\_and\\_architecting\\_for\\_machine\\_learning.pdf](https://www.gartner.com/binaries/content/assets/events/keywords/catalyst/catus8/preparing_and_architecting_for_machine_learning.pdf).

**SYSTEMIC: Information System and Informatics Journal**

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 1 - Agustus 2020

**Analisis Perbandingan Sensitivitas Metode Simple Additive Weighting dan Weighted Product pada Pemilihan Kos di Sekitar UIN Sunan Kalijaga Yogyakarta**Irsalina Santi Khasanah<sup>1</sup>, Agus Mulyanto<sup>2</sup>, Muhammad Galih Wonoseto<sup>3</sup><sup>1,2,3</sup> UIN Sunan Kalijaga Yogyakarta[15650008@student.uin-suka.ac.id](mailto:15650008@student.uin-suka.ac.id)<sup>1</sup>, [agus.mulyanto@uin-suka.ac.id](mailto:agus.mulyanto@uin-suka.ac.id)<sup>2</sup>, [muhhammad.wonoseto@uin-suka.ac.id](mailto:muhhammad.wonoseto@uin-suka.ac.id)<sup>3</sup>**Kata Kunci***Sensitivitas, Simple Additive Weighting, Weighted Product***Abstrak**

Metode dalam sistem pendukung keputusan sangat beragam sehingga kita perlu mengetahui metode mana yang lebih sensitif dalam penyelesaian sebuah kasus. Penelitian ini bertujuan untuk melakukan perbandingan sensitivitas pada metode Simple Additive Weighting dan metode Weighted Product dalam pemilihan kos di sekitar UIN Sunan Kalijaga Yogyakarta. Analisis sensitivitas dilakukan pada 6 kali percobaan dengan mengubah nilai bobot mulai dari 1, 1.5, dan 2 dengan menaikkan nilai bobot sebesar 0.5 dan 0.1. Berdasarkan penelitian yang dilakukan, sensitivitas pada Metode SAW adalah 31.26% dan sensitivitas pada Metode WP adalah 30.53%. Pada penelitian ini metode SAW memiliki nilai sensitivitas yang lebih tinggi dibandingkan dengan WP.

**Keywords***Sensitivity, Simple Additive Weighting, Weighted Product***Abstract**

The methods in decision support systems are so diverse that we need to know which methods are more sensitive in solving a case. This study aims to compare the sensitivity of the Simple Additive Weighting method and the Weighted Product method in the selection of boarding around UIN Sunan Kalijaga Yogyakarta. Sensitivity analysis was carried out on 6 trials by changing the weight values starting from 1, 1.5, and 2 by increasing the weight values by 0.5 and 0.1. Based on research conducted, the sensitivity in the SAW Method is 31.26% and the sensitivity in the WP Method is 30.53%. In this study, the SAW method has a higher sensitivity value compared to WP.

**1. Pendahuluan**

Kos digunakan untuk memenuhi kebutuhan hunian yang bersifat sementara dengan sasaran pada umumnya adalah pelajar yang berasal dari luar daerah. Banyak faktor yang dipertimbangkan oleh calon penghuni kos dalam memilih kos, misalnya harga, fasilitas, luas kamar, dan lain sebagainya.

UIN Sunan Kalijaga Yogyakarta merupakan perguruan tinggi negeri yang memiliki letak strategis dan pilihan program studi yang beragam, mengakibatkan meningkatnya minat masyarakat untuk melanjutkan studi di UIN Sunan Kalijaga Yogyakarta.

Meningkatnya jumlah mahasiswa UIN Sunan Kalijaga Yogyakarta mempengaruhi munculnya banyak kos yang terdapat di sekitar kampus. Hal ini mengakibatkan mahasiswa baru kesulitan untuk memilih kos. Permasalahan tersebut dapat diatasi menggunakan Sistem Pendukung

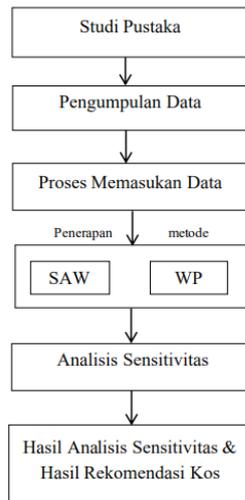
Keputusan, yang merupakan sistem berbasis komputer yang interaktif dalam membantu pengambilan keputusan memanfaatkan data dan model untuk menyelesaikan suatu masalah [1].

Metode yang digunakan peneliti dalam SPK ini adalah Simple Additive Weighting (SAW) dan Weighted Product (WP). Metode Simple Additive Weighting (SAW) bekerja dengan cara mencari penjumlahan terbobot dari rating kinerja pada alternatif pada semua kriteria atau atribut. Sedangkan metode Weighted Product (WP) bekerja dengan melakukan perbandingan dengan menggunakan perkalian untuk menghubungkan rating atribut [2]. Pada penelitian ini peneliti akan membandingkan sensitivitas kedua metode tersebut.

**2. Metode Penelitian**

Dalam penelitian ini, terdapat beberapa tahapan yang akan dilakukan. Tahap-tahapan

tersebut dijelaskan pada gambar 1 berikut ini:



Gambar 1. Metode Penelitian

## 2.1 Studi Pustaka

Studi Pustaka dilakukan dengan membaca dan mempelajari referensi bacaan yang berupa jurnal ilmiah, skripsi, tesis, serta buku. Studi pustaka juga dilakukan dengan media internet melalui pencarian data-data atau informasi yang berkaitan dengan objek penelitian seperti sistem pendukung keputusan, metode Weighted Product (WP) dan Analisis sensitivitas. Pengumpulan Data.

Peneliti melakukan pengumpulan data dengan wawancara dan kuesioner. Wawancara dilakukan melalui tanya jawab dengan beberapa mahasiswa UIN Sunan Kalijaga. Untuk kuesioner, terdapat dua jenis yaitu kuesioner yang menentukan kriteria dan kuesioner yang menentukan data kos di sekitar UIN Sunan Kalijaga Yogyakarta. Kuesioner dibuat untuk mengetahui jawaban responden atas beberapa pertanyaan tertutup untuk kuesioner kriteria dan pertanyaan terbuka untuk kuesioner data Kos. Kuesioner pertama untuk menentukan kriteria disusun menggunakan skala likert. Biasanya dalam skala ini rentang penilaian dari 1 sampai 5 dengan beberapa format skala [3] yaitu:

- 1) Skor 1= Sangat tidak X.
- 2) Skor 2= Tidak X.
- 3) Skor 3= Cukup X.
- 4) Skor 4= X.
- 5) Skor 5= Sangat X.

Pertanyaan dalam kuesioner pertama berupa kriteria yang dapat menjadi pertimbangan dalam memilih kos di sekitar UIN Sunan Kalijaga. Kuesioner kedua peneliti membuat daftar pertanyaan untuk mendapatkan informasi data kos di sekitar UIN Sunan Kalijaga.

Dalam pengumpulan data, digunakan juga rumus slovin [4] dalam penentuan jumlah sample alternative yang akan dipakai dalam penelitian ini.

$$n = \frac{N}{1 + N(e)^2} \quad (1)$$

Dengan

$n$  = ukuran sample

$N$  = Populasi

$e$  = Persentase kelonggaran ketelitian yang dapat ditolerir  $e = 0.1$  (10%).

## 2.2 Proses Memasukkan Data

Data yang dimasukkan dalam sistem ini berupa:

- 1) Kriteria yang dipilih
- 2) Penambahan bobot untuk melakukan analisis sensitivitas

## 2.3 Penerapan Metode SAW dan WP

Proses penerapan metode Simple Additive Weighting (SAW) dan Weighted Product (WP) dilakukan setelah adanya proses memasukkan data. Dalam proses ini akan menghasilkan rangking perhitungan kedua metode tersebut yang nantinya dijadikan basis mengetahui awal dalam proses analisis sensitivitas.

### Penerapan Metode SAW

Proses dari metode SAW terdiri dari beberapa langkah [2]:

- 1) Mendefinisikan terlebih dahulu kriteria-kriteria yang akan menjadi tolak ukur penyelesaian masalah.
- 2) Menormalisasi setiap nilai alternatif pada setiap atribut rating
- 3) Menghitung nilai bobot preferensi pada setiap alternatif.
- 4) Melakukan perbandingan

### Penerapan Metode WP

- 1) Mendefinisikan terlebih dahulu kriteria-kriteria yang akan di jadikan sebagai tolak ukur penyelesaian masalah.
- 2) Menormalisasi setiap nilai alternatif (nilai vector).
- 3) Menghitung nilai bobot preferensi pada setiap alternatif.
- 4) Melakukan perbandingan.

### Hasil SAW dan WP

Setelah melakukan input kriteria dan proses perhitungan maka akan di dapat hasil rangking metode Simple Additive Weighting (SAW) dan Weighted Product (WP) yang nantinya akan dijadikan patokan dalam perubahan rangking dalam proses analisis sensitivitas.

## 2.4 Analisis Sensitivitas

Tahap penelitian selanjutnya adalah proses analisis sensitivitas. Analisa sensitivitas dapat dikatakan sebagai suatu cara untuk mengetahui pengaruh pada solusi optimal yang dihasilkan oleh metode simpleks jika parameter diubah nilainya. Manfaat utama dari analisa sensitivitas adalah untuk mengidentifikasi parameter yang sensitif [5].

Proses analisis sensitivitas dimulai dengan melakukan penambahan bobot dalam setiap

implementasi metode pengambilan keputusan untuk melihat derajat sensitivitasnya. Analisis sensitivitas menggunakan derajat sensitivitas yang telah ditentukan sebagai parameter analisis.

**2.5 Hasil Analisis Sensitivitas dan Hasil Rekomendasi Kos**

Dalam pemilihan metode dengan analisis sensitivitas adalah semakin sensitif hasil perangkaan setiap metode maka metode tersebut akan semakin dipilih. Hasil akhir pemilihan kos adalah rangking nilai alternatif. Semakin tinggi nilai alternatif, maka alternatif tersebut merupakan alternatif terbaik.

**3. Hasil Dan Pembahasan**

**3.1 Proses Pengambilan Data**

Data yang digunakan dalam penelitian ini adalah data kos di sekitar UIN Sunan Kalijaga Yogyakarta sebanyak 41 data dan data kriteria yang didapatkan melalui kuesioner. Setelah mendapatkan data maka dapat digunakan sebagai penentuan data alternatif, data kriteria dan data subkriteria.

**Penentuan Data Alternatif**

Penentuan data alternatif, merupakan menentukan data kos yang berada di sekitar UIN Sunan Kalijaga Yogyakarta. Data alternatif diambil sejak tanggal 14 Januari – 31 Januari 2019 melalui kuesioner google form. Dan dalam penelitian ini menggunakan data sampel dengan mencari jumlah data kos yang berada di sekitar UIN Sunan Kalijaga pada website <https://mamikos.com/kost/kost-dekatuin-jogja-murah>, yaitu didapatkan sebanyak 69 data kos di sekitar UIN Sunan Kalijaga dan dilakukan perhitungan mencari data sampel. Untuk mencari jumlah data sampel maka dilakukan perhitungan dengan menggunakan rumus slovin seperti rumus (1) diatas. Dengan N = 69, maka dilakukan perhitungan sebagai berikut:

$$\frac{69}{1 + 69x(0.1)^2} = 40,82 \text{ (dibulatkan menjadi 41)}$$

Dari perhitungan diatas, didapat nilai n=41, sehingga jumlah sample yang digunakan adalah 41. Keempat puluh satu alternatif kos tersebut adalah:

Table 1. Alternatif

Kode	Nama Kos	Kode	Nama Kos
A1	Wisma Al-Maun	A21	Bugenvil
A2	Wisma Gading 2	A22	Kost Ladies
A3	Kos Multazam	A23	Azka
A4	Kos concat	A24	Wisma Arundina
A5	Kos ledok tukang	A25	Kos Putri
A6	Kos pinggir kali	A26	Wisma Toples

A7	Kos Pak Suciyono	A27	Kos Siswati 8A
A8	Kos pak tanto	A28	Kos mas emen
A9	Kos GGS Ganteng Ganteng Sholeh	A29	Kost Putri Karunia
A10	Wisma BengQeng	A30	Kos babe soebarman
A11	Kos 482	A31	Wisma Eidelwis
A12	Kos Tan Panama	A32	Elite kost
A13	Kos Gendeng	A33	Kos Kasmaran
A14	Kos 820A Ambarukmo	A34	Kos Bu Jevi
A15	Putri Muslim Almira	A35	Kos Annisa
A16	Larasati	A36	Kos TPI
A17	Mikandry	A37	Kos Bu Handri
A18	Hibrida 1	A38	Kos kasmaran
A19	Kos Hijau Muslimah	A39	Kos bu tuti
A20	Kos Sabila	A40	Kos Istana Salon
		A41	Kos Ibu mar

**Penentuan Data Kriteria**

Terdapat tujuh kriteria yang dijadikan acuan dalam pengambilan keputusan adapun penentuan data kriteria, ditentukan berdasarkan hasil survei yang telah dilakukan sebelumnya menggunakan google form. Adapun tujuh kriteria yang telah ditentukan adalah: harga, fasilitas, luas kamar, jarak, akses lingkungan, kebersihan dan keamanan.

Tahap selanjutnya adalah menentukan nilai bobot kriteria. Penentuan prioritas dan bobot kriteria ditentukan melalui data hasil kuesioner online dengan google form, yang telah diisi oleh 148 responden. namun dalam menentukan bobot kriteria peneliti hanya menggunakan 130 data yaitu responden yang pernah mencari kos dan yang pernah kebingungan dalam mencari kos.

Penentuan nilai bobot hanya dilakukan pada kriteria K1 hingga K7. Penentuan nilai responden menggunakan skala likert dengan ketentuan nilai 1, 2, 3, 4, dan 5. Dan telah dinormalisasi oleh peneliti maka didapatkan hasil normalisasi bobot kriteria yang disajikan pada tabel berikut ini:

Table 2. Kriteria

Kode	Kriteria	Tipe	Bobot
K1	Akses Lingkungan	Benefit	0.21
K2	Luas	Benefit	0.18
K3	Jarak	Cost	0.15
K4	Fasilitas	Benefit	0.14
K5	Harga	Cost	0.12
K6	Kebersihan	Benefit	0.10
K7	Keamanan	Benefit	0.10

**Penentuan Sub kriteria**

Data penilaian didapat dari masing-masing kriteria, lalu dibuat suatu tingkat kepentingan kriteria berdasarkan nilai bobot yang telah ditentukan dalam skala tingkat kepentingan. Rating kecocokan setiap alternatif pada setiap

kriteria adalah sebagai berikut:

- Sangat Rendah (SR) = 0.1
- Rendah (R) = 0.2
- Cukup (C) = 0.3
- Tinggi (T) = 0.4
- Sangat Tinggi (ST) = 0.5

Setelah itu, langkah selanjutnya adalah menentukan interval (rentang skala) untuk masing-masing kelas. Untuk mencari interval kelas maka dilakukan perhitungan dengan menggunakan Rumus Struges [6].

$$CI = \frac{Range(R)}{C} \quad (2)$$

Dengan:

- CI = Interval Kelas
- R = Range
- C = Banyak Kelas

Dari rumus tersebut, kita terapkan pada semua sub kriteria:

- 1) Range sub kriteria keamanan kos: Dengan skala linkert pada kuesioner dengan interval yang dikehendaki 0.1 - 0.5 (sangat tidak aman, tidak aman, cukup aman, aman, sangat aman)
- 2) Range sub kriteria kebersihan kos: Dengan skala linkert pada kuesioner dengan interval yang dikehendaki 0.1 - 0.5 (sangat tidak bersih, tidak bersih, cukup bersih, bersih, sangat bersih)
- 3) Range sub kriteria harga kos:

$$\frac{12jt - 2.3jt}{5} = 1.94jt \text{ (dibulatkan jadi 2 jt)}$$

- 4) Range sub kriteria fasilitas kos:

$$\frac{11 - 2}{5} = 1.8$$

- 5) Range sub kriteria jarak kos:

$$\frac{2000 - 100}{5} = 380 \text{ (dibulatkan jd 400 m)}$$

- 6) Range sub kriteria luas kos:

$$\frac{20 - 6}{5} = 2.8m^2 \text{ (dibulatkan jd } 3m^2)$$

Table 3. Sub Kriteria

Kriteria	Sub Kriteria	Range	Nilai Bobot
Keamanan	Sangat Tidak Aman	Sangat Tidak Aman	0.1
	Tidak Aman	Tidak Aman	0.2
	Cukup aman	Cukup aman	0.3
	Aman	Aman	0.4
	Sangat Aman	Sangat Aman	0.5

Kebersihan	Sangat Tidak Bersih	Sangat Tidak Bersih	0.1
	Tidak Bersih	Tidak Bersih	0.2
	Cukup Bersih	Cukup Bersih	0.3
	Bersih	Bersih	0.4
	Sangat Bersih	Sangat Bersih	0.5
Harga	Sangat Murah	2.1 jt - 4 jt	0.1
	Murah	4.1 jt - 6 jt	0.2
	Cukup Murah	6.1 jt - 8 jt	0.3
	Mahal	8.1 jt - 10 jt	0.4
	Sangat Mahal	10.1 jt - 12 jt	0.5
Fasilitas	Sangat Tidak Lengkap	2-3 fasilitas	0.1
	Kurang Lengkap	4-5 fasilitas	0.2
	Cukup Lengkap	6-7 fasilitas	0.3
	Lengkap	8-9 fasilitas	0.4
	Sangat Lengkap	10-11 fasilitas	0.5
Jarak	Sangat Jauh	1601-2000m	0.5
	Jauh	1201-1600m	0.4
	Cukup Jauh	801-1200m	0.3
	Dekat	401-800m	0.2
	Sangat Dekat	1-400m	0.1
Akses Lingkungan	Sangat Tidak Lengkap	2-3 akses	0.1
	Tidak Lengkap	4-5 akses	0.2
	Cukup Lengkap	6-7 akses	0.3
	Lengkap	8-9 akses	0.4
	Sangat Lengkap	10-11 akses	0.5
Luas	Tidak Luas	6-8m	0.1
	Kurang Luas	9-11m	0.2
	Cukup Luas	12-14m	0.3
	Luas	15-17m	0.4
	Sangat Luas	18-20m	0.5

### 3.2 Proses Memasukkan Data

#### 3.2.1 Kriteria yang dipilih

Kriteria yang dipilih adalah semua kriteria yang ada pada table 2.

#### 3.2.2 Penambahan bobot untuk melakukan analisis sensitifitas

Tahap Input penambahan bobot dalam analisis sensitivitas peneliti melakukan enam kali percobaan dalam menambahkan bobot kriteria.

- 1) Percobaan 1: Mengubah bobot kriteria dengan menambahkan bobot sebesar 1 dengan proses penambahan bobot di mulai dari 0.5
- 2) Percobaan 2: Mengubah bobot kriteria dengan menambahkan bobot sebesar 1.5 dengan proses penambahan bobot di mulai dari 0.5
- 3) Percobaan 3: Mengubah bobot kriteria dengan menambahkan bobot sebesar 2 dengan proses penambahan bobot di mulai dari 0.5
- 4) Percobaan 4: Mengubah bobot kriteria dengan

- menambahkan bobot sebesar 1 dengan proses penambahan bobot di mulai dari 0.1
- 5) Percobaan 5: Mengubah bobot kriteria dengan menambahkan bobot sebesar 1.5 dengan proses penambahan bobot di mulai dari 0.1
- 6) Percobaan 6: Mengubah bobot kriteria dengan menambahkan bobot sebesar 2 dengan proses penambahan bobot di mulai dari 0.1

**3.3 Proses Perhitungan SAW dan WP**

Sebelum melakukan analisis sensitivitas, kita perlu melakukan perhitungan metode Simple Additive Weighting (SAW) dan Weighted Product (WP) terlebih dahulu untuk mendapatkan basis pengetahuan awal untuk dilakukan tahap analisis sensitivitas.

**3.3.1 Perhitungan SAW**

Pada tahap ini, dilakukan tahap-tahap metode SAW dalam pemilihan kos. Diantaranya [2]:

- 1) Input Kriteria  
Tahap ini adalah menginputkan ketujuh kriteria yang sebelumnya sudah kita jabarkan pada table 2.
- 2) Membuat Matrix Keputusan  
Terdapat dua tahap dalam membuat matrix keputusan.
  - Mendefinisikan matrix keputusan yang berukuran m x n, dengan m= jumlah alternative dan n=jumlah kriteria
  - Memberikan nilai pada setiap alternative dan setiap kriteria pada matrix keputusan X.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} \\ x_{i1} & x_{i2} & \dots & x_{ij} \end{bmatrix} \quad (3)$$

- Normalisasi Matrix Keputusan  
Melakukan normalisasi matriks keputusan Z dengan cara menghitung nilai rating kinerja ternormalisasi (rij) dari alternatif Ai pada atribut Cj ditunjukkan pada rumus berikut ini:

$$r_{ij} = \begin{cases} \frac{x_{ij}}{\max_i x_{ij}} & \text{jika } j \text{ adalah atribut keuntungan (benefit)} \\ \frac{\min_i x_{ij}}{x_{ij}} & \text{jika } j \text{ adalah atribut biaya (cost)} \end{cases} \quad (4)$$

Hasil normalisasi matrix keputusan tidak dapat ditampilkan dalam paper ini mengingat ukurannya yang begitu besar.

- 3) Melakukan proses pembobotan
  - Melakukan proses perangkingan dengan cara mengalikan matrik ternormalisasi (N) dengan nilai bobot preferensi (W).
  - Menentukan nilai preferensi untuk setiap alternatif (Vi) dengan cara menjumlahkan hasil kali antara matrik ternormalisasi (N) dengan nilai bobot preferensi (W) yang ditunjukkan pada rumus:

$$V_i = \sum_{j=1}^n w_j r_{ij} \quad (5)$$

- 4) Hasil Rekomendasi Kos berdasarkan Metode SAW  
Setelah dilakukan beberapa tahap perhitungan maka hasil rekomendasi kos terbaik metode Simple Additive Weighting (SAW) adalah kos Wisma BengQeng, dengan beberapa urutan ranking sebagai berikut:

Table 4. Hasil Rekomendasi Kos dengan Metode SAW

Ranking	Kode	Nama Kos	Jumlah
1	A10	Wisma BengQeng	0,8260
2	A22	Kost Ladies	0,7780
3	A31	Wisma Eidelwis	0,7560
4	A6	Kos pinggir kali	0,7500

Dst...

**3.3.2 Perhitungan WP**

Pada tahap ini, dilakukan tahap-tahap metode WP dalam pemilihan kos, diantaranya:

- 1) Input Kriteria  
Proses input kriteria dilakukan dengan memasukan prioritas kriteria yang sudah ditentukan oleh peneliti, dengan memasukan urutan prioritas kriteria yang di dapat melalui kuesioner.
- 2) Mencari Vektor S
  - Menentukan range untuk penentuan bobot kriteria.
  - Memasukan bobot preferensi tiap kriteria sebagai tolak ukur perhitungan.
  - Menentukan kriteria Cost dan Benefit.
  - Menghitung preferensi alternatif.

Preferensi alternatif yang dianalogikan dengan vektor S. Vektor S dapat dihitung dengan cara mengangkat dan mengalikan nilai masing-masing kriteria dengan bobot yang sudah diperbaiki. Vektor S dapat dihitung dengan rumus:

$$S_i = \prod_{j=1}^n X_{ij}^{w_j} ; \text{ dengan } i = 1,2,.. (6)$$

Dengan:

- S= Preferensi alternatif dianalogikan sebagai vektor S
- x = nilai kriteria
- w = bobot kriteria
- i = alternative
- j = kriteria
- m = banyaknya kriteria

- 3) Menghitung Preferensi Alternatif Vektor

Menghitung Preferensi relatif dari setiap alternatif yang diberikan diberikan. Nilai preferensi relatif dapat dihitung dengan rumus:

$$V_i = \frac{\prod_{j=1}^n x_{ij} w_j}{\prod_{j=1}^n (x_{j*}) w_j}; \text{ dengan } i = 1, 2, \dots, m \quad (7)$$

Dengan:

V = Preferensi alternatif dianalogikan sebagai vektor

Vx = nilai kriteria

w = bobot kriteria

i = alternative

j = kriteria

m = banyaknya kriteria

\* = banyaknya kriteria yang telah dinilai pada vektor S

- 4) Hasil Rekomendasi Kos berdasarkan Metode WP

Setelah melakukan tahap-tahap perhitungan metode Weighted Product (WP) maka di dapatkan hasil rekomendasi kos terbaik adalah kos Wisma BengQeng dengan urutan ranking sebagai berikut:

Table 5. Rekomendasi Kos Metode WP

Ranking	Kode	Nama Kos	Jumlah
1	V10	Wisma BengQeng	0,034583
2	V22	Kost Ladies	0,031748
3	V31	Wisma Eidelwis	0,030759
4	V21	Bugenvil	0,029873

Dst...

### 3.4 Proses Analisis Sensitivitas

Tahap ini bertujuan untuk mendapatkan perubahan ranking alternatif dalam pemilihan kos disekitar UIN Sunan Kalijaga Yogyakarta. Semakin besar perubahan ranking alternative, maka semakin sensitive metode tersebut. Menurut Yeh [7] yang dikutip dalam paper yang ditulis oleh Rachmat [8], berikut ini merupakan langkah-langkah perhitungan analisis sensitivitas dalam kasus pengambilan keputusan, yaitu:

- Tentukan semua bobot atribut,  $w_j = 1$  ( $w$  = bobot awal kriteria), dengan  $j = 1, 2, \dots, n$  jumlah atribut.
- Mengubah bobot salah satu kriteria sedangkan bobot kriteria yang lain bernilai tetap. (Ubah bobot suatu atribut dengan menambah nilai bobot dalam range 1 - 2, misalkan dengan menaikkan nilai bobot sebesar 0.1. Sementara bobot atribut lainnya masih bernilai tetap. Penambahan 0.1 tidak baku dan tidak ada ketentuan mengikat.
- Normalisasi bobot atribut tersebut dengan cara membentuk nilai bobot sedemikian

hingga  $\Sigma = 1$ .

- Aplikasikan metode MADM yang akan digunakan, untuk bobot-bobot atribut yang telah dibentuk.
- Hal yang harus diperhatikan disini adalah proses analisis sensitivitas dilakukan dalam waktu bersamaan kesetiap metode yang akan dianalisis. Jadi ketika perulangan pertama selesai dilakukan maka setiap metode tersebut akan menyimpan nilai atau hasil perankingan sementara.
- Hitung persentase perubahan ranking dengan cara membandingkan berapa banyak perubahan ranking yang terjadi jika dibandingkan dengan kondisi pada saat bobot sebelum dilakukan analisis sensitivitas. Nilai persentase perubahan perankingan alternatif dapat diperoleh dengan rumus:

$$\frac{T}{i \times A} \times 100\% \quad (8)$$

Dengan:

T = total akhir perubahan ranking alternatif pada setiap metode MADM.

I = total perulangan.

A = atribut yang digunakan pada setiap kasus.

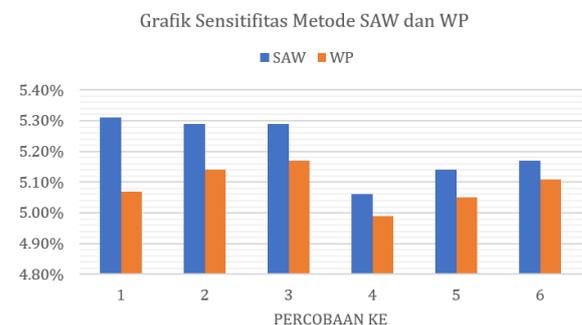
### 3.5 Hasil Percobaan

Proses analisis sensitivitas pada penelitian ini menggunakan 6 percobaan dalam penambahan bobot kriteria. Berikut ini hasil percobaannya:

coba ke	Total Iterasi	Perubahan Ranking		Persentase Sensitivitas	
		SAW	WP	SAW	WP
1	14	520	479	5.31%	5.07%
2	21	778	755	5.29%	5.14%
3	28	1036	1014	5.29%	5.17%
4	70	2480	2444	5.06%	4.99%
5	105	3775	3715	5.14%	5.05%
6	140	5056	5008	5.17%	5.11%
Jumlah	378	13654	13433	31.26%	30.53%

#### 3.5.1 Hasil Percobaan Sensitivitas

Setelah dilakukan enam kali percobaan, didapatkan hasil seperti pada grafik berikut ini:



#### 4 Kesimpulan

Berdasarkan penelitian yang telah dilakukan didapatkan kesimpulan bahwa analisis perbandingan sensitivitas metode Simple Additive Weighting (SAW) dan metode Weighted Product (WP) bisa digunakan untuk menentukan metode yang relevan dalam pemilihan kos di sekitar UIN Sunan Kalijaga Yogyakarta.

Proses analisis sensitivitas pada 6 percobaan dengan menambah bobot yang berbeda menyatakan bahwa Metode Simple Additive Weighting (SAW) memiliki dominasi nilai persentase sensitivitas yang lebih tinggi dibandingkan dengan Metode Weighted Product (WP).

Analisis sensitivitas pada Metode SAW dengan 6 percobaan menghasilkan jumlah perubahan ranking sebesar 13654 dan jumlah persentase sensitivitas sebesar 31.26%, sedangkan analisis sensitivitas pada Metode WP dengan 6 percobaan menghasilkan jumlah perubahan ranking sebesar 13344 dan jumlah persentase sensitivitas sebesar 30.53%.

Metode SAW memiliki nilai perubahan ranking dan persentase sensitivitas yang lebih tinggi dibandingkan dengan WP, sehingga pengambilan keputusan yang relevan dalam pemilihan kos disekitar UIN Sunan Kalijaga Yogyakarta dapat dilakukan dengan Metode SAW.

*Int. Trans. Oper. Res.*, vol. 9, no. 2, pp. 169–181, 2002.

- [8] A. G. Rachmat, B. Suprpty, and A. Najib, “Perbandingan Metode SAW dan Topsis Untuk Penentuan Dosen Terbaik Pada Jurusan teknologi Informasi Politeknik Negeri Samarinda,” *Pros. Semin. Ilmu Komput. dan Teknol. Inf.*, vol. 3, no. 1, pp. 91–97, 2018, [Online]. Available: <http://e-journals.unmul.ac.id/index.php/SAKTI/article/view/2073/pdf>.

#### Daftar Pustaka

- [1] E. Turban, J. E. Aronson, and T. P. Liang, *Decision Support Systems and Intelligent Systems*, 7 th. New Jerseyfer: Pearson Education, Inc, 2005.
- [2] S. Kusumadewi, S. Hartati, A. Harjoko, and R. Wardoyo, *Fuzzy Multi-Attribute Decision Making (Fuzzy MADM)*, Edisi Pert. Yogyakarta: Penerbit Graha Ilmu, 2006.
- [3] D. Nofriansyah and S. Defit, *Buku Multi Criteria Decision Making (MCDM) pada Sistem Pendukung Keputusan*. Yogyakarta: CV Budi Utama, 2017.
- [4] H. Umar, *Metode Riset Bisnis: Panduan Mahasiswa untuk Melaksanakan Riset Dilengkapi Contoh Proposal dan Hasil Riset Bidang Manajemen dan Akuntansi*, 2nd ed. Jakarta: PT. Gramedia Pustaka Utama, 2003.
- [5] H. W. Subandi and T. B. Adji, *Analisis Sensitivitas untuk Pemilihan Model Multi Attribute Decision Making (MADM) dalam Penyelesaian Masalah Dinamis*. Yogyakarta: Universitas Gadjah Mada (UGM), 2012.
- [6] P. B. Santosa and M. Hamdani, *Statistika Deskriptif dalam Bidang Ekonomi dan Niaga*. Semarang: PT Gelora Aksara Pratama, 2007.
- [7] C. Yeh, “A Problem-based Selection of Multi-attribute Decision-making Methods,”

**SYSTEMIC: Information System and Informatics Journal**

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 1 – Agustus 2020

**Sistem Pembelajaran Hukum Baca Al-Qur'an Menggunakan Algoritma LPC dan KNN**Hafizh Achmad Dinan<sup>1</sup>, Youllia Indrawaty N<sup>2</sup>, Kurnia Ramadhan Putra<sup>3</sup><sup>1,2,3</sup>) Institut Teknologi Nasional[hafizhdinan35@gmail.com](mailto:hafizhdinan35@gmail.com)<sup>1</sup>, [youllia1975@gmail.com](mailto:youllia1975@gmail.com)<sup>2</sup>, [kurniaramadhanp91@gmail.com](mailto:kurniaramadhanp91@gmail.com)<sup>3</sup>**Kata Kunci***LPC, KNN, ekstraksi suara, python, tajwid.***Abstrak**

Seorang muslim harus dapat membaca ayat-ayat Al-Quran dengan baik sesuai yang diajarkan oleh Rasulullah saw. Membaca Al-Quran sesuai ilmu tajwid hukumnya wajib bagi setiap orang muslim, apabila seseorang membaca Al-Quran dengan tidak memakai tajwid, hukumnya berdosa. Pengembangan Aplikasi Pembelajaran Tajwid Al-Qur'an ini bertujuan untuk membantu seorang muslim agar pandai membaca al-Qur'an yang baik dan benar. Surat yang di pakai dalam aplikasi ini adalah surat Al-Fatihah. Aplikasi Pembelajaran Tajwid Al-Qur'an menggunakan Bahasa pemrograman Python dengan metode Linear Predictive Coding (LPC) sebagai ekstraksi ciri suara dan metode K-Nearest Neighbor sebagai pencocokan dengan data latih. Untuk pengujian pengucapan ayat ke-1 didapatkan akurasi data sebesar 83.3%, ayat ke-2 sebesar 86.7%, ayat ke-3 sebesar 85%, ayat ke-4 sebesar 80%, ayat ke-5 sebesar 88.3%, ayat ke-6 sebesar 93.3%

**Keywords***LPC, KNN, python, , tajwid voice extraction.***Abstract**

A Muslim must be able to read the verses of the Qur'an properly as taught by the Prophet, Muhammad. Reading the Qur'an in accordance with tajwid is obligatory for every Muslim, if someone reads the Qur'an without using tajwid, the law is sinful. The development of the application of learning the Tajwid of Qur'an is aimed at helping a Muslim to be good at reading the Qur'an that is good and right. Al-Fatihah is uses in this application. Learning the Tajwid of Qur'an Application is using Linear Predictive Coding (LPC) method as sound feature extraction and K-Nearest Neighbor as matching with training data. For testing the pronunciation of the 1st verse obtained data accuracy of 83.3%, the 2nd verse is 86.7%, the 3rd verse is 85%, the 4th verse is 80%, the 5th verse is 88.3%, the 6th verse is 93.3%.

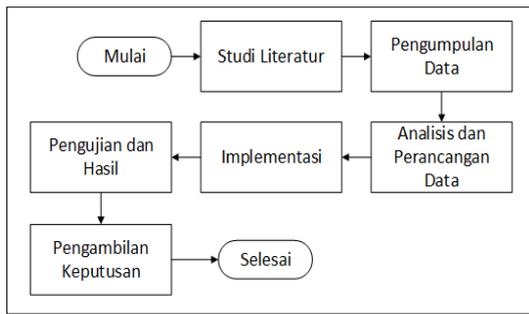
**1. Pendahuluan**

Seorang muslim harus dapat membaca ayat-ayat Al-Quran dengan baik sesuai yang diajarkan oleh Rasulullah saw. Membaca Al-Quran sesuai ilmu tajwid hukumnya wajib bagi setiap orang muslim, tidak bisa diwakili oleh orang lain. Apabila seseorang membaca Al-Quran dengan tidak memakai tajwid, hukumnya berdosa karena sesungguhnya Allah SWT menurunkan Al-Quran berikut tajwidnya[1]. Salah membaca hukum bacaan Al-Qur'an dapat menimbulkan arti yang berbeda, meskipun itu hanya panjang pendeknya suara. Surat Al-Fatihah merupakan surat pertama dari kitab suci Al-Qur'an, dan selalu diucapkan saat shalat.

Berdasarkan permasalahan tersebut, yaitu saat membaca Al-Qur'an orang masih sering salah dalam membaca tajwid Al-Qur'an. Oleh karena itu, penelitian ini fokus kepada merancang aplikasi mendeteksi suara baca Al-Qur'an yang dibangun dengan Algoritma LPC untuk ekstraksi ciri suara dan Algoritma KNN untuk pencocokan dengan *database*.

**2. Metode Penelitian**

Beberapa tahapan penelitian yang dilakukan meliputi studi literatur dan pengumpulan data, analisis dan perancangan sistem, implementasi sistem, pengujian dan hasil, serta kesimpulan. Alur dari tahapan penelitian tersebut dapat dilihat pada Gambar 1.



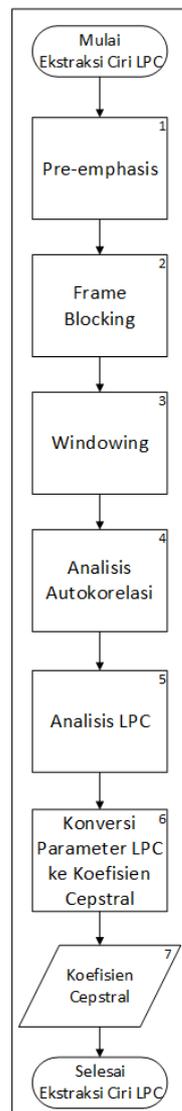
Gambar 1. Alur Metode Penelitian

## 2.1 Pengumpulan Data

Data penelitian yang digunakan merupakan hasil rekaman dan dari situs yang beralamat <https://islamdownload.net/> dalam bentuk file .wav.

## 2.2 Linear Predictive Coding (LPC)

Terdapat beberapa penelitian yang menggunakan metode LPC karena menurut Rabiner [2] LPC memiliki kelebihan yaitu sederhana dan dapat diterapkan pada perangkat lunak dan keras.



Gambar 2. Langkah-langkah LPC

### a. Pre-emphasis

Tahap pertama adalah menghilangkan noise suara yang tidak dibutuhkan. Noise ini berupa sinyal yang memiliki perbedaan signifikan di antara nilai di sekitarnya. Pada tahapan ini diterapkan Persamaan (1) (Deng, O'Shaughnessy, 2003):

$$y(n) = s(n) - \alpha \cdot s(n - 1) \quad (1)$$

$\tilde{s}(n)$  : sinyal suara setelah dilakukan proses preemphasize.

$y(n)$  : sinyal suara sebelum dilakukan proses preemphasize.

$\tilde{\alpha}$  : nilai yang paling sering digunakan yaitu 0.95

### b. Frame Blocking

Frame Blocking merupakan proses membagi sinyal suara menjadi beberapa blok suara.

### c. Windowing

Pada tahapan frame blocking ada kemungkinan terdapat sinyal yang terputus-putus atau mengalami diskontinuitas antara blok satu dengan yang lainnya, sehingga pada tahapan windowing dilakukan untuk mengurangi efek tersebut. Dengan menerapkan Persamaan (3) [2] dengan asumsi jumlah sample masing-masing frame atau N adalah 640, sehingga diperoleh  $w(n)$  sebagai berikut:

$$w(n) = 0,54 - 0,46 \cos\left(\frac{2\pi n}{N-1}\right) \quad (2)$$

$$\tilde{x}(n) = x(n) \cdot w(n) \quad (3)$$

$W(n)$  adalah nilai *hamming*

$\tilde{x}(n)$  adalah nilai *windowing*

$x(n)$  adalah nilai *frame blocking*

### d. Analisis Autokorelasi

Bagian sinyal yang telah melalui tahap windowing, dibentuk autokorelasinya dengan Persamaan (4) berikut:

$$r(m) = \sum_{n=0}^{N-1-m} \tilde{x}(n) \cdot \tilde{x}(n+m) \quad (4)$$

$r(m)$  adalah nilai analisis autokorelasi

$\tilde{x}(n)$  adalah nilai *windowing*

### e. Analisis LPC

Proses ini dilakukan untuk mengubah nilai autokorelasi menjadi parameter koefisien transformasi. Pada tahap ini akan menghasilkan sinyal suara melalui perhitungan berikut:

$$E^{(0)} = r(0) \quad (5)$$

$$E^{(i)} = (1 - k_i^2) E^{(i-1)} \quad (6)$$

$$k_i = \frac{r(i) - \sum_{j=1}^{i-1} \alpha_j^{(i-1)} r(i-j)}{E^{(i-1)}} \quad (7)$$

$$\alpha_i^{(i)} = k_i \quad (8)$$

$$\alpha_j^{(i)} = \alpha_j^{(i-1)} - k_i \alpha_{i-j}^{(i-1)} \quad (9)$$

E adalah nilai error

K adalah koefisien PARCOR

$\alpha$  adalah koefisien LPC

f. Konversi Parameter LPC ke Koefisien Cepstral  
 Koefisien cepstral ini adalah koefisien dari representasi transformasi Fourier pada spectrum logaritmis

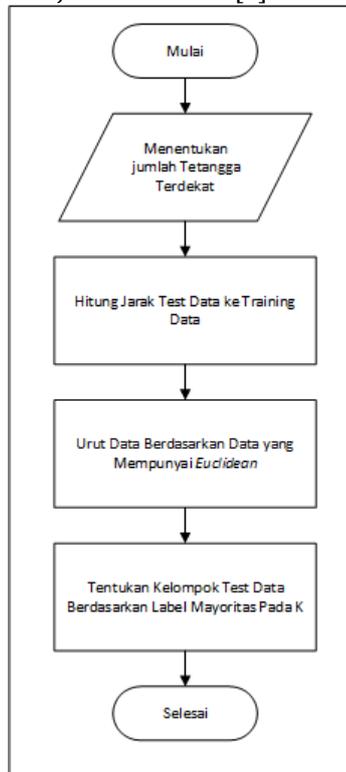
$$c_m = \alpha_m + \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) \cdot c_k \cdot \alpha_{m-k} \quad (10)$$

$c$  adalah koefisien Cepstral  
 $a$  adalah koefisien LPC  
 $m$  adalah orde ke- $n$

**2.3 K-Nearest Neighbor**

Dalam pengenalan pola, algoritma tetangga- $k$  terdekat ( $k$ -NN) adalah metode non-parametrik yang digunakan untuk klasifikasi dan regresi.[3]

Algoritma K-Nearest Neighbor mengelompokkan data baru yang belum diketahui class-nya dengan memilih data sejumlah  $k$  yang letaknya paling dekat dari data baru. Class paling banyak dari data terdekat sejumlah  $k$  akan dipilih sebagai class yang diprediksi untuk data baru [4]. Pada umumnya nilai  $k$  adalah jumlah ganjil supaya tidak terdapat jarak yang sama dalam proses klasifikasi. Jauh atau dekatnya tetangga dihitung menggunakan jarak Euclidean[5].



Gambar 3. Flowchart Metode KNN

Pada Gambar 3 dapat dilihat bahwa tahapan metode KNN melalui beberapa tahapan berikut:

1. Tahap awal adalah menentukan tetangga terdekat ( $k$ ).
2. Selanjutnya menghitung dengan rumus euclidean distance antara sample latihan dan sample uji dengan Persamaan (10).

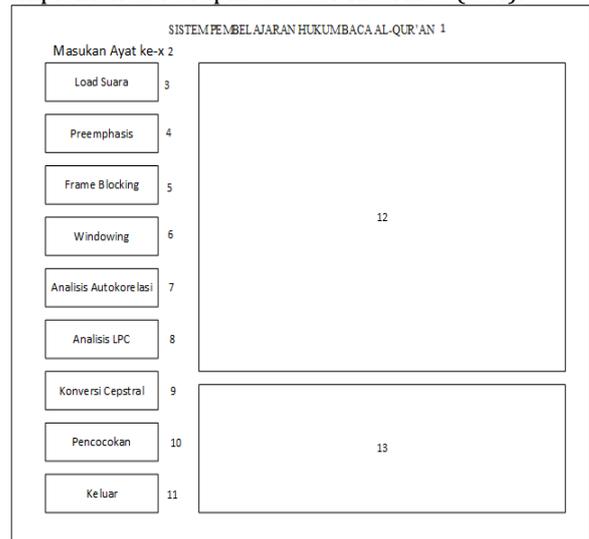
$$D(a, b) = \sqrt{\sum_{k=1}^d (a_k - b_k)^2} \quad (11)$$

- $D(a,b)$  : Nilai Euclidean Distance
- $a$  : Data latihan (training)
- $b$  : Data uji (testing)
- $k$  :  $k$  (fitur) ke-
- $d$  : Dimensi (jumlah  $k$ )

3. Setelah itu mengurutkan data yang mempunyai nilai  $k$ .
4. Tahap akhir mengelompokkan sample latihan berdasarkan mayoritas nilai  $k$ .

**2.4 Implementasi**

Implementasi pada penelitian ini terdiri dari implementasi Graphical User Interface (GUI).



Gambar 4. User Interface

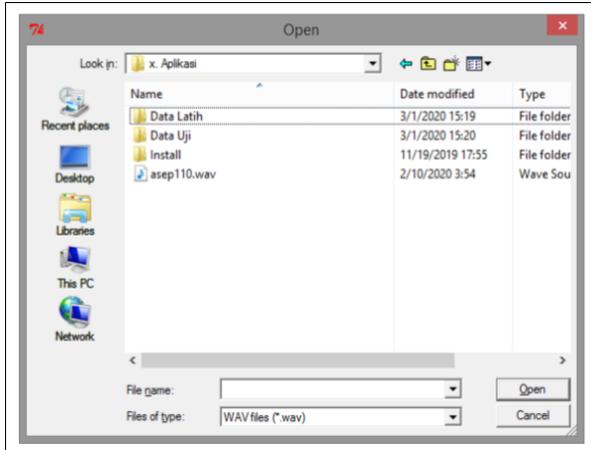
Keterangan dari Gambar 4 sebagai berikut:

1. Judul Aplikasi
2. Indikator Ayat ke-x
3. Tombol load suara
4. Tombol preemphasis
5. Tombol Frame Blocking
6. Tombol Windowing
7. Tombol Analisis Autokorelasi
8. Tombol Analisis LPC
9. Tombol Konversi Cepstral
10. Tombol Pencocokan
11. Tombol Keluar
12. Grafik
13. Textbox

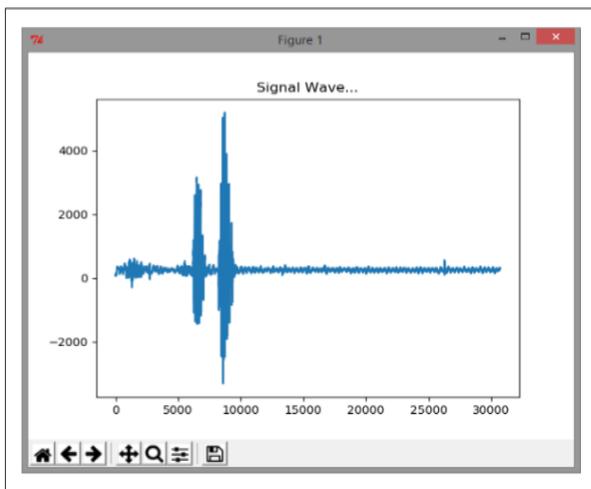
**3. Hasil Dan Pembahasan**

Sebelum dilakukan proses pencocokkan, data uji dan data latihan diekstraksi dahulu menggunakan Linear Prediction Coding untuk mendapatkan nilai koefisien cepstral suara masukan. Pada penelitian penentuan kualitas suara ini terdapat tujuh pengucapan ayat surat Al-Fatihah. Dimulai dengan proses pengambilan data dengan cara ambil data suara yang sudah ada. Pengguna merekam atau mengambil data suara yang berformat .wav dengan tombol Browse suara yang berformat .wav.

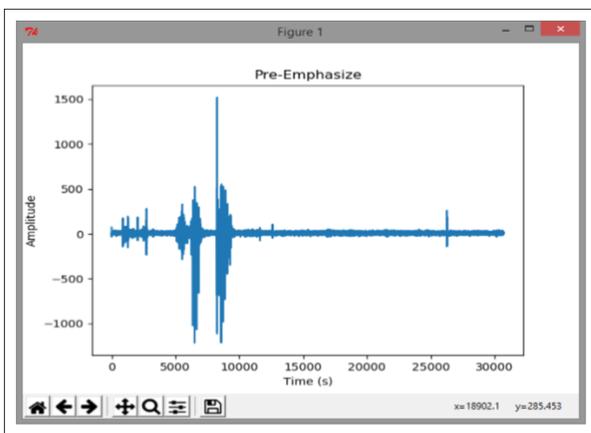
Gambar 5 menunjukkan pengguna mengambil data suara berformat .wav yang sudah ada sebelumnya dan suara yang diambil akan dijadikan data suara uji untuk penelitian ini. Gambar 6 merupakan tampilan sinyal dari data suara yang dipilih ditampilkan pada layar aplikasi. Dan Gambar 7 merupakan hasil dari data sinyal suara yang sudah dilakukan proses *Preemphasize*.



Gambar 5. Browse Suara



Gambar 6. Hasil Sinyal Suara



Gambar 7. Hasil Preemphasis

Data sinyal suara yang telah diambil akan dilakukan proses pertama yaitu *Preemphasize* yang bertujuan untuk mendapatkan data sinyal suara

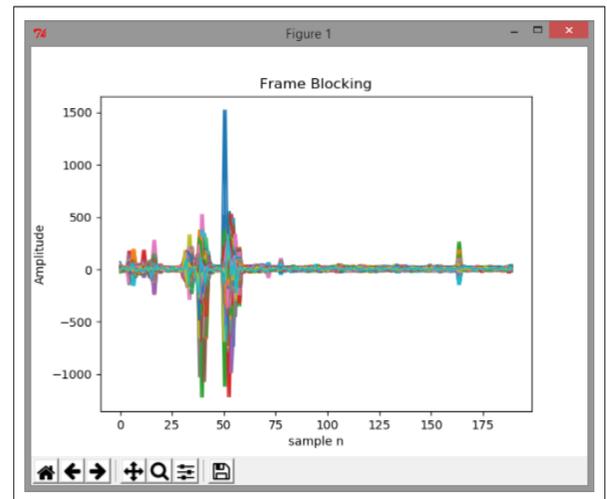
tanpa noise.

Misalkan nilai adalah  $(-20.125, -36.125, -25.325, -4.425, 19.175, 25.275, 42.875, 27.875)$ , sampel sinyal :  $(-20.125, -36.125, -25.325, -4.425, 19.175, 25.275, 42.875, 27.875)$ , dengan  $\alpha = 0.95$  dengan menerapkan Persamaan (1), maka diperoleh

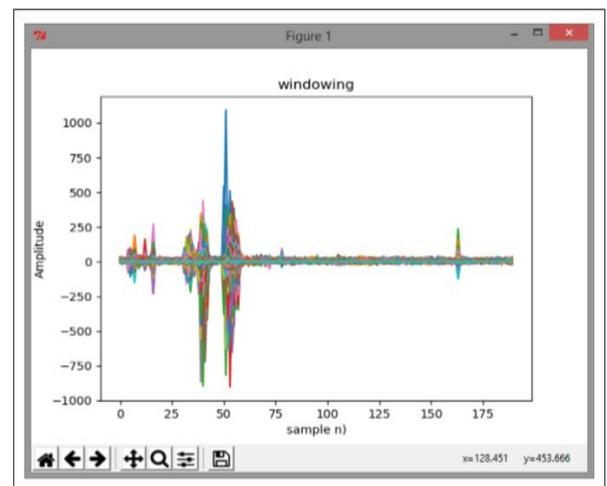
$$\begin{aligned} y(0) &= s(0) = -20.125 \\ y(1) &= s(1) - \alpha * y(1 - 1) \\ &= -36.125 - (-20.125 * 0.95) \\ &= -17.006 \end{aligned}$$

Begitu juga dengan ketujuh data yang lainnya, sehingga didapatkan nilai sinyal setelah preemphasis adalah  $(-20.125, -17.006, 8.994, 19.634, 23.379, 7.059, 18.864, -12.856)$ .

Gambar 8 merupakan hasil dari data sinyal suara yang sudah dilakukan proses *Frame Blocking*. Hasil data sinyal suara yang sudah diproses pada tahapan *Preemphasize* sinyal akan disegmentasi menjadi beberapa frame ditahapan *Frame Blocking* ini.



Gambar 8. Hasil Frame Blocking



Gambar 9. Hasil Windowing

Gambar 9 merupakan hasil dari data sinyal suara yang sudah dilakukan proses *Windowing*. Hasil data sinyal suara yang sudah disegmentasi pada tahapan *Frame Blocking* akan menimbulkan efek diskontinuitas di awal dan di akhir frame. Saat

dilakukan proses windowing, data sinyal suara yang memiliki diskontinuitas di awal dan di akhir diubah nilainya menjadi nol.

Diketahui nilai sinyal setelah Frame Blocking adalah : (1.824 , 1.1771, 1.1664, 1.2551, 0.8692, 0.0165, -0,7363, -0.0334) dengan nilai N = 320 dari perhitungan N pada bagian frame blocking sebelumnya. Pertama mencari nilai *hamming* dengan persamaan (2):

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right)$$

$$w(0) = 0.54 - 0.46\cos\left(\frac{2 * 3.14 * 0}{320 - 1}\right)$$

$$w(0) = 0.54 - 0.46\cos(0)$$

$$w(0) = 0.54 - 0.46 * 1$$

$$w(0) = 0.54 - 0.46 = 0.08$$

Kemudian mencari nilai windowing dengan persamaan (3):

$$\tilde{x}(0) = x(0) \cdot w(0)$$

$$\tilde{x}(0) = 1.824 * 0.08 = 0.14592$$

Begitu juga dengan ketujuh data yang lainnya, sehingga didapatkan nilai sinyal setelah *windowing* yaitu (0.14592 , 0.094168, 0.093312 ,0.100408, 0.069536 , 0.00132, - 0.058936 , -0.002672)

Selanjutnya tahap Analisis Autokorelasi yaitu korelasi silang terhadap sinyal dengan dirinya sendiri.

Jika data hasil windowing sampel data latih adalah (0.192 , 0.168, 0.312 ,0. 408, 0.536 , 0.32, - 0.936 , -0.672) maka nilai autokorelasinya dengan menggunakan persamaan (4) adalah:

$$r(0) = \tilde{x}(0) * \tilde{x}(0) + \tilde{x}(1) * \tilde{x}(1) + \tilde{x}(2) * \tilde{x}(2) + \tilde{x}(3) * \tilde{x}(3) + \tilde{x}(4) * \tilde{x}(4) + \tilde{x}(5) * \tilde{x}(5) + \tilde{x}(6) * \tilde{x}(6) + \tilde{x}(7) * \tilde{x}(7)$$

$$r(0) = (0.192 * 0.192) + (0.168 * 0.168) + (0.312 * 0.312) + (0.408 * 0.408) + (0.536 * 0.536) + (0.32 * 0.32) + (-0.936 * (-0.936)) + (-0.672 * (-0.672))$$

$$r(0) = 2.046272$$

Begitu juga dengan ke tujuh data lainnya sehingga didapatkan hasil analisis autokorelasi menggunakan order 8 adalah (2.046, 0.932, -0.290, -0.474, -0.410, -0.305, -0.293, -0.129).

Selanjutnya tahap Analisis LPC yaitu untuk mengubah nilai autokorelasi menjadi parameter koefisien transformasi dengan menggunakan persamaan (5), (6), (7), (8),(9):

$$E^{(0)} = 2.046$$

$$k_1 = \frac{r(1) - \sum_{j=1}^{1-1} \alpha_j^{(1-1)} r(|1-j|)}{E^{(1-1)}}$$

$$k_1 = \frac{r(1)}{E^{(0)}} = \frac{0.932}{2.046} = 0.45552297165 = 0.456$$

$$\alpha_1^{(1)} = k_1$$

$$\alpha_1^{(1)} = 0.456$$

$$E^{(1)} = (1 - k_1^2)E^{(1-1)}$$

$$E^{(1)} = (1 - (0.456)^2)2.046 = 1.620562944 = -1.621$$

Begitu juga dengan ke tujuh data lainnya sehingga didapatkan hasil analisis LPC adalah (1.654, 1.415, 0.522, -0.578, -1.146, -1.641, 1.057)

Tahap terakhir dalam metode LPC adalah Konversi koefisien LPC menjadi koefisien Cepstral dengan menggunakan persamaan (10):

$$c_0 = \ln(p) = \ln(7) = 1.94591014906 = 1.946$$

$$c_1 = \alpha_1 + \sum_{k=1}^{1-1} \left(\frac{k}{1}\right) \cdot c_k \cdot a_{1-k} = a_1 = 1.645$$

$$c_2 = \alpha_2 + \sum_{k=1}^{2-1} \left(\frac{k}{2}\right) \cdot c_k \cdot a_{2-k} = \alpha_2 + \left(\frac{1}{2} * c_1 * \alpha_1\right)$$

Begitu juga dengan data lainnya sehingga didapatkan hasil koefisien Cepstral adalah (1.94591, -0.01799, -0.12958, 0.10819, -0.19576, -0.41561, 0.06887, -0.11295). Kemudian, setelah tahapan mendapatkan nilai cepstral, dilakukan pencocokan dengan metode KNN.

Tabel 1 merupakan nilai cepstral dari 6 file suara wav terdiri dari 3 file suara ayat 1 dan 3 file suara ayat 2. Nilai koefisien cepstral yang akan di ui adalah C0 = 1.946 , C1= -0.304, C2= -0.058, C3= 0.109, C4= -0.375, C5=0.105, C6=0.261, C7=0.067. Perhitungan KNN ada dalam Tabel 2 menggunakan persamaan (11). Setelah dihitung *square distance* nya, Selanjutnya mengurutkan jarak tersebut dan tentukan tetangga mana yang terdekat berdasarkan jarak minimum ke -K sebagaimana Tabel 3. Selanjutnya menentukan kategori dari tetangga terdekat. Kategori dari tetangga terdekat (Y) tidak termasuk karena rangking dari data ini lebih dari 3 (=K). terdapat kategori dari yaitu satu dan dua karena di database hanya terdapat klasifikasi satu dan dua sebagaimana Tabel 4.

Tabel 1. Cepstral dan Klasifikasi

C0 Data Latih	C1 Data Latih	C2 Data Latih	C3 Data Latih	C4 Data Latih	C5 Data Latih	C6 Data Latih	C7 Data Latih	Klasifikasi Data Latih (Ayat)
1.946	-0.018	-0.13	0.108	-0.196	-0.416	0.069	-0.113	1
1.946	0.099	0.182	-0.619	-0.474	0.009	0.061	0.183	1
1.946	-0.131	0.043	0.244	-0.549	0.11	-0.091	0.11	1
1.946	0.041	0.512	0.235	0.679	-0.194	-0.18	0.211	2
1.946	-0.047	0.724	-0.8	-0.113	0.01	-0.04	0.016	2
1.946	0.063	-0.203	-0.281	-0.541	0.59	0.085	0.251	2

Tabel 2. Square Distance

Square Distance
$(1.946 - 1.946)^2 + (-0.018 - (-0.304))^2 + (-0.13 - (-0.058))^2 + (0.108 - 0.109)^2 + (-0.196 - (-0.375))^2 + (-0.416 - 0.105)^2 + (0.069 - 0.261)^2 + (-0.113 - 0.067)^2 = 0.459727$
$(1.946 - 1.946)^2 + (0.099 - (-0.304))^2 + (0.182 - (-0.058))^2 + (-0.619 - 0.109)^2 + (-0.474 - (-0.375))^2 + (0.009 - 0.105)^2 + (0.061 - 0.261)^2 + (0.183 - 0.067)^2 = 0.822466$
$(1.946 - 1.946)^2 + (-0.131 - (-0.304))^2 + (0.043 - (-0.058))^2 + (0.244 - 0.109)^2 + (-0.549 - (-0.375))^2 + (0.11 - 0.105)^2 + (-0.091 - 0.261)^2 + (0.11 - 0.067)^2 = 0.214409$
$(1.946 - 1.946)^2 + (0.041 - (-0.304))^2 + (0.152 - (-0.058))^2 + (0.235 - 0.109)^2 + (0.679 - (-0.375))^2 + (-0.194 - 0.105)^2 + (-0.18 - 0.261)^2 + (0.211 - 0.067)^2 = 1.594535$
$(1.946 - 1.946)^2 + (-0.047 - (-0.304))^2 + (0.724 - (-0.058))^2 + (-0.8 - 0.109)^2 + (-0.113 - (-0.375))^2 + (0.1 - 0.105)^2 + (-0.04 - 0.261)^2 + (0.016 - 0.067)^2 = 1.665725$
$(1.946 - 1.946)^2 + (0.063 - (-0.304))^2 + (-0.203 - (-0.058))^2 + (-0.281 - 0.109)^2 + (-0.541 - (-0.375))^2 + (0.59 - 0.105)^2 + (0.085 - 0.261)^2 + (0.251 - 0.067)^2 = 0.635427$

Tabel 3. Mengurutkan data berdasarkan jarak minimum-K

Square Distance	Urutan Jarak	Apakah termasuk kategori k=3 ? (rangking 1,2 dan 3)
0.459727	2	Ya
0.822466	4	Tidak
0.214409	1	Ya
1.594535	5	Tidak
1.665725	6	Tidak
0.635427	2	Ya

Tabel 4. Menentukan Kategori

Square Distance	Urutan Jarak	Apakah termasuk kategori k=3 ? (rangking 1,2 dan 3)	Y = Kategori dari KNN (Class yang terdapat pada Data Latih)
0.459727	2	Ya	Ayat Satu
0.822466	4	Tidak	-
0.214409	1	Ya	Ayat Satu
1.594535	5	Tidak	-
1.665725	6	Tidak	-
0.635427	2	Ya	Ayat Dua

Selanjutnya digunakan kategori mayoritas yang sederhana dari tetangga yang terdekat tersebut sebagai nilai prediksi dari data yang baru. Dari data tersebut di dapatkan dua kategori angka 1 dan satu kategori angka 2, karena kategori angka 1 lebih banyak maka data koefisien cepstral uji yang di masukkan adalah termasuk dalam kategori satu atau ayat 1.

Hasil dari pencocokan KNN seperti pada Gambar 10 jika benar dan jika pencocokan knn salah maka seperti pada Gambar 13. jika ayatnya benar tetapi bacaan madnya terlalu panjang atau pendek akan muncul pesan seperti Gambar 11 dan Gambar 12 dan menambah total salah.

```
Prediksi =['satu']
Benar, Suara ini Ayat ke-1
```

Gambar 10. Pencocokan KNN Benar

```
Prediksi =['satu']
Benar, Suara ini Ayat ke-1, tetapi bacaannya terlalu pendek
```

Gambar 11. Pencocokan Benar Bacaan Pendek

```
Prediksi =['satu']
Benar, Suara ini Ayat ke-1, tetapi bacaannya terlalu panjang
```

Gambar 12. Pencocokan Benar Bacaan Panjang

```
Prediksi =['satu']
Salah, Suara ini bukan Ayat ke-2
```

Gambar 13. Pencocokan KNN Salah

Jika user melakukan kesalahan, maka counter total salah akan bertambah 1 seperti pada Gambar 14.

Total Salah = 1 x

Gambar 14. Total Salah

Setelah semua ayat tekeh di uji oleh aplikas, akan muncul berapa kali *user* melakukan kesalahan pada ayat ke-n tersebut. Hasil akhir program seperti pada Gambar 15.

Ayat 1 Salah = 0 x
Ayat 2 Salah = 2 x
Ayat 3 Salah = 0 x
Ayat 4 Salah = 1 x
Ayat 5 Salah = 0 x
Ayat 6 Salah = 4 x

Gambar 15. Hasil Akhir Program

Hasil perhitungan ekstraksi ciri LPC pada 5 orang sampel berada pada Tabel 5. Sedangkan Tabel 6 adalah contoh pengujian KNN pada orang ke-1.

Tabel 5. Nilai Koefisien Cepstral

Orang ke-	Ayat ke-					
	1	2	3	4	5	6
1	1.94591	1.94591	1.94591	1.94591	1.94591	1.94591
	-0.01799	0.04074	0.13428	0.00618	0.0409	0.00404
	-0.12958	0.51162	1.17207	0.36951	0.11517	-0.20692
	0.10819	0.2346	0.31266	0.1153	0.04061	-0.42255
	-0.19576	0.67899	-0.20616	-0.15279	0.43143	0.42252
	-0.41561	-0.19398	0.04201	0.08586	0.06945	-0.09284
	0.06887	-0.17948	-0.22339	-0.12729	0.052	0.04057
	-0.11295	0.21143	-0.04667	0.10328	-0.0798	0.09415
	2	1.94591	1.94591	1.94591	1.94591	1.94591
0.0999		-0.04718	0.13613	0.07253	-0.04409	0.0215
0.1822		0.72355	1.44149	1.1112	-0.03485	0.56928
-0.6192		-0.79969	-0.78311	-0.12762	0.49943	-0.04506
-0.47412		-0.1129	0.29446	0.18066	-0.34606	-0.28338
0.00861		0.00967	-0.19702	-0.22659	0.06432	-0.20474
0.06136		-0.03951	0.21343	0.02258	0.13542	-0.10947
0.18327		0.01588	0.13034	0.13908	0.02815	-0.08204
3		1.94591	1.94591	1.94591	1.94591	1.94591
	-0.13134	0.06323	0.22712	0.08494	0.07202	0.0526
	0.0426	-0.20343	-0.4161	0.41539	0.76058	-0.67125
	0.24386	-0.28123	-0.35287	0.22398	-0.12786	-0.27071
	-0.54857	-0.54047	0.09688	-0.08801	0.18567	-0.09328
	0.11005	0.05895	-0.15741	0.23639	-0.44394	0.05121
	-0.09054	0.08528	0.35055	0.0903	-0.19414	0.12524
	0.11042	0.25059	-0.16364	-0.05854	0.14731	0.04725
	4	1.94591	1.94591	1.94591	1.94591	1.94591
-0.14684		0.14846	0.24895	0.14806	0.15863	-0.12284
0.26266		-0.25471	0.08419	1.54646	1.53458	-0.17955
0.1467		-0.18829	-0.03937	-0.80319	-0.79396	-0.24963
0.11007		-0.06999	0.18201	0.36836	0.35265	-0.21152
-0.18807		0.0119	0.02719	-0.25703	-0.24802	-0.11265
-0.09203		0.53442	0.0655	0.2561	0.24993	0.415
0.51436		-0.1714	0.02945	0.13789	0.14275	-0.04152
5		1.94591	1.94591	1.94591	1.94591	1.94591
	0.16758	0.17228	-0.2593	0.16496	0.16839	0.16938
	1.51104	1.39253	0.89231	0.20365	0.39622	1.47274
	-0.77235	-0.7871	0.13424	-0.00554	0.04305	-0.75962
	0.30698	0.13404	-0.25098	-0.30007	-0.25721	0.266
	-0.22149	-0.23708	-0.07297	-0.11607	0.35844	-0.20197
	0.22203	0.11949	0.038	0.07304	-0.18693	0.19826
	0.16163	0.14803	-0.01332	-0.22912	0.06786	0.16128

Tabel 6. Pencocokan KNN pada orang ke-1

Ayat	C0	C1	C2	C3	C4	C5	C6	C7	
1	1.946	-0.018	-0.13	0.108	-0.196	-0.416	0.069	-0.113	Berhasil
2	1.946	0.041	0.512	0.235	0.679	-0.194	-0.18	0.211	Berhasil
3	1.946	0.134	1.172	0.313	-0.206	0.042	-0.223	-0.047	Gagal
4	1.946	0.006	0.37	0.115	-0.153	0.086	-0.127	0.103	Gagal

5	1.946	0.041	0.115	0.041	0.431	0.069	0.052	-0.08	Berhasil
6	1.946	0.004	-0.207	-0.423	0.423	-0.093	0.041	0.094	Berhasil

Persamaan (12), (13), dan (14) adalah persamaan untuk menghitung *precision*, *recall*, dan *accuracy* pada setiap ayat.

$$precision = \frac{TP}{(TP+FP)} \times 100\% \quad (12)$$

$$recall = \frac{TP}{(TP+FN)} \times 100\% \quad (13)$$

$$accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)} \times 100\% \quad (14)$$

TP : True Positive, jumlah ayat ke-n yang benar yang terdeteksi oleh aplikasi

TN : True Negative, jumlah ayat ke-n yang terdeteksi benar oleh aplikasi tetapi bukan ayat ke-n

FP : False Positive, jumlah suara ayat ke-n yang benar tetapi tidak terdeteksi oleh aplikasi

FN : False Negative, jumlah suara yang bukan ayat ke-n

Tabel 7. Data Pengujian

Ayat ke-	TP	FP	FN	TN
1	6	6	4	44
2	5	3	5	47
3	6	5	4	45
4	4	6	6	44
5	6	3	4	47
6	8	2	2	48

Tabel 3 adalah data pengujian dari 60 file suara wav dengan setiap ayat 10 file suara wav.

Berikut adalah contoh perhitungan dengan persamaan (11), (12), (13) menggunakan contoh ayat ke-1:

$$Precision \text{ ayat 1} = \frac{TP}{(TP + FP)} \times 100\% = \frac{6}{6 + 6} \times 100\% = 50\%$$

$$Recall \text{ ayat 1} = \frac{TP}{(TP + FN)} \times 100\% = \frac{6}{6 + 4} \times 100\% = 60\%$$

$$Accuracy \text{ ayat 1} = \frac{(TP+TN)}{(TP+FP+FN+TN)} \times 100\% = \frac{6 + 44}{(6 + 6 + 4 + 44)} \times 100\% = 83.3\%$$

Setelah dihitung menggunakan persamaan (12), (13), dan (14) maka didapatkan hasil seperti pada Tabel 8.

Tabel 8 Akurasi Data

Ayat ke-	Precision	Recall	Accuracy
1	50%	60%	83.3%
2	62.5%	50%	86.7%
3	54.6%	60%	85%
4	40%	40%	80%
5	66.7%	60%	88.3%
6	80%	80%	93.3%

#### 4. Kesimpulan

Berdasarkan penelitian mengenai sistem ukum baca Al-Qur'an yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

Untuk mendapatkan nilai koefisien cepstral dilakukan dengan cara pencarian metode Linear Prediction Coding dan pencocokan antara data suara Koefisien Cepstral latih dengan data suara Koefisien Cepstral uji menggunakan metode K-nearest Neighbor (KNN). Sistem yang dibangun dapat digunakan untuk mengenali suara berupa ayat.

Untuk pengujian pengucapan ayat ke-1 didapatkan akurasi data sebesar 83.3%, ayat ke-2 sebesar 86.7%, ayat ke-3 sebesar 85%, ayat ke-4 sebesar 80%, ayat ke-5 sebesar 88.3%, ayat ke-6 sebesar 93.3%.

#### Daftar Pustaka

- [1] Suryani, D., Irfan, M., Uriawan, W., & Budiawan, W. (2016). IMPLEMENTASI ALGORITMA DIVIDE AND CONQUER PADA APLIKASI BELAJAR ILMU TAJWID. JOIN Volume I No. 1.
- [2] Rabiner, L., & Juang, B.-H. (1933). Fundamental of Speech Recognition. New Jersey: PTR Prentice-Hall, Inc.
- [3] Altman, N. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. American Statistical Association, 175-185.
- [4] Shofia, E. N., Putri, R. R., & Arwan, A. (2017). Sistem Pakar Diagnosis Penyakit Demam: DBD, Malaria dan Tifoid Menggunakan Metode K-Nearest Neighbor - Certainty Factor. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 426-435.
- [5] Han, J., Kamber, M., & Pei, J. (2012). *Data Mining. Concept and Techniques*. Waltham: Morgan Kaufmann.
- [6] Dewi, I. A., Zulkarnain, A., & Lestari, A. A. (2018). Identifikasi Suara Tangisan Bayi menggunakan Metode LPC dan Euclidean Distance. *ELKOMIKA, Vol. 6, No. 1*, 153-164.
- [7] Fatoni, C. S., & Noviandha, F. D. (2017). Case Based Reasoning Diagnosis Penyakit Difteri dengan Algoritma K-Nearest Neighbor. *Citec Journal*, 220-232.
- [8] Hanafi, M., Fadillah, N., & Ihsan, A. (2019). Optimasi Algoritma K-Nearest Neighbor untuk Klasifikasi Tingkat Kematangan Buah Alpukat Berdasarkan Warna. *IT Journal Research and Development (ITJRD)*, 10-18.
- [9] S Harsemadi, G., Sudarma, M., & Pramaita, N.

- (2017). Implementasi Algoritma K-Nearest Neighbor pada Perangkat Lunak Pengelompokan Musik untuk Menentukan Suasana Hati. *Teknologi Elektro*, 15-19.
- [10] Idwal, A. Y., Nurhasanah, Y. I., & Utami, D. B. (2017). Sistem Pengenalan Suara Bahasa Indonesia Untuk Mengenali Aksen Melayu Pontianak Dan Sunda Garut. *Jurnal Teknik Informatika dan Sistem Informasi Volume 3 Nomor 3*, 461-471.
- [11] Indrawaty, Y., Dewi, I. A., & Lukman, R. (2019). Ekstraksi Ciri Pelafalan Huruf Hijaiyyah Dengan Metode Mel-Frequency Cepstral Coefficients. *MIND Journal*, 01-16.
- [12] Mustakim, & Oktaviani F, G. (2016). Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa. *Jurnal Sains, Teknologi dan Industri*, 195-202
- [13] Nurhasanah, Y. I., Dewi, I. A., & Saputro, B. A. (2018). Iqro Reading Learning System through Speech Recognition Using Mel Frequency Cepstral Coefficient (MFCC) and Vector Quantization (VQ) Method. *Int. Journal of Applied IT Vol. 02 No. 01*, 29-42.
- [14] Rivki, M., & Bachtiar, M. (2017). Implementasi Algoritma K-Nearest Neighbor Dalam Pengklasifikasian Follower Twitter Yang Menggunakan Bahasa Indonesia. *Jurnal Sistem Informasi (Journal of Information Systems)*, 31-37.
- [15] Sudaradjat, D. (2019). Digitalisasi Sinyal Suara Manusia Dengan Algoritma Linear Predictive Coding. *Jurnal Ilmu Pengetahuan Dan Teknologi Komputer*, 177-184
- [16] A Triansyah, E., & Indrawaty N, Y. (2017). Implementasi Metode Pattern Recognition Untuk Pengenalan Ucapan Huruf Hijaiyyah. *Jurnal Ilmiah Teknologi Informasi Terapan*, 1-10.

# SYSTEMIC: Information System and Informatics Journal

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 1 – Agustus 2020

## Sistem Penunjang Keputusan Seleksi Atlet Berdasarkan Data Fisik Menggunakan Naïve Bayes

Yosia Christmas Decky Halundaka<sup>1</sup>, Din Syamsudin<sup>2</sup>, Aryo Nugroho<sup>3</sup><sup>1,2,3</sup> Universitas Narotama, Surabaya[yosia.chrismas@mhs.fasilkom.narotama.ac.id](mailto:yosia.chrismas@mhs.fasilkom.narotama.ac.id)<sup>1</sup>, [din.syamsudin@fik.narotama.ac.id](mailto:din.syamsudin@fik.narotama.ac.id)<sup>2</sup>, [aryo.nugroho@narotama.ac.id](mailto:aryo.nugroho@narotama.ac.id)<sup>3</sup>

### Kata Kunci

Atlet, Klasifikasi, Naïve Bayes, Sistem Pendukung Keputusan, Tes Fisik.

### Abstrak

Tujuan dari penelitian ini adalah untuk memprediksi tingkat kelulusan hasil test fisik peserta atlet beladiri kempo yang diselenggarakan oleh KONI. Peneliti melakukan klasifikasi berdasarkan data yang diperoleh dari pelatih fisik beladiri kempo dengan menggunakan data mining. Data mining dapat diartikan sebagai proses pengekstrakan informasi baru yang diambil dari bongkahan data besar untuk menunjang hasil pengambilan Sistem Pendukung Keputusan. Oleh karena itu menggunakan Sistem Pendukung Keputusan yang dapat membantu pengambilan keputusan untuk memilih atlet terbaik. Dalam data mining terdapat banyak metode yang dapat digunakan, salah satunya adalah metode klasifikasi dengan memanfaatkan algoritma naive bayes. Hasil yang diperoleh dari penelitian ini adalah Correctly Classified Instances 94.4% dengan speed presentase 60% dan 70% menghasilkan Correctly Classified Instances 92,85%, sedangkan speed presentase 80% menghasilkan Correctly Classified Instances 88,89%.

### Keywords

Athlete, Classification, Decision Support System, Fisically Test, Naive Bayes

### Abstract

The purpose of this observation was to predict the passing rate of the physical test results of the participants of the martial arts athlete organized by KONI. The researchist made a classification based on data obtained from trainer of physical martial arts Kempo by using data mining. Data mining can be interpreted is a process extractcing new informartion was taken from big chunk of data for support the results of making decision support systems. Therefor using decision support systems which can help make a decision to choose the best athlete. In data mining there are many methods that can be used, one of them is the classification method by using alogarithms naive bayes. The results obtained from this observation is Correctly Classified Instances 94.4% with speed presentase 60% and 70% bring out Correctly Classified Instances 92.85%, while speed presentase 80% bring out Correctly Classified Instances 88.89%.

## 1. Pendahuluan

Pengelompokan nilai data fisik atlet beladiri menggunakan Metode Naïve Bayes merupakan salah satu elemen dasar yang harus ada dalam setiap atlet beladiri, karena pengelompokan nilai data fisik atlet beladiri sangat berpengaruh dalam proses pengembangan atlet tersebut, salah satu tolak ukur penguasaan materi pengajaran oleh atlet. Data nilai tes fisik perlu dikelompokkan untuk membedakan nilai yang baik dan buruk dengan jangkauan kelompok nilai tertentu. Hasil pengelompokan nilai ini dapat digunakan untuk membuat suatu kebijakan kemajuan fisik atlet menggunakan metode Naïve Bayes. Kebijakan tersebut dapat berupa pembenahan metode pembenahan fisik atlet, atau

pengukuran tingkat keberhasilan atlet dalam melakukan tes fisik[1].

Sebelum tes dimulai atlet terlebih dahulu harus melakukan *stretching* (peregangan) dan *warming up* (pemanasan) setelah melakukan pemanasan tes fisik pertama akan di mulai yaitu tes *sprint* yang dipantau langsung oleh ahli fisik asal Australia Mr. Greg . Kemudian dilanjutkan tes berikutnya. selama tes fisik berlangsung atlet tidak diberi waktu jeda untuk beristirahat. dari semua tes fisik yang ada, tes fisik VO2Max adalah tes fisik yang paling sukar[2]. Karena hanya beberapa orang saja yang mampu menyelesaikan tes fisik VO2Max [3]. Setelah semua tes berjalan dengan lancar maka berikutnya melakukan tes BIA (*Bioelectrical Impedance Analysis*) untuk

mengetahui masa otot masa lemak dan kadar air seorang atlet dan memperkirakan kadar otot lemak atlet cukup ideal atau tidak karena kecukupan otot lemak dan air menentukan peak performa atlet ketika bertanding. Jika kadar tersebut merasa kurang atau lebih maka asupan atlet akan dikhususkan [4][5].

Salah satu penerapan pengelompokan data adalah Naïve Bayes. Naïve Bayes adalah algoritma klasifikasi yang paling sederhana dibanding algoritma klastering yang lain. Algoritma ini mempunyai kelebihan mudah diterapkan dan dijalankan, relatif cepat, mudah untuk adaptasi, dan paling banyak dipraktekkan dalam tugas data mining. Algoritma ini termasuk salah satu algoritma paling penting dalam data mining. Diharapkan dengan adanya metode ini dapat membantu meningkatkan fisik atlet bela diri.

Aturan dan metode Naïve Bayes dapat diterapkan pada sebuah program bantu untuk mengelompokkan data nilai seorang atlet beladiri. Langkah-langkah diterjemahkan ke dalam bahasa pemrograman untuk melakukan tugas klasifikasi data. Dataset nilai atlet dimasukkan ke dalam input program, kemudian program melakukan pengolahan sesuai langkah algoritma Naive Bayes, dan hasilnya berupa klasifikasi data. Data inilah yang digunakan sebagai bahan pertimbangan untuk pembuatan kebijakan pada atlet.

**2. Metode Penelitian**

Tujuan penelitian ini adalah untuk mengetahui dan mengelompokkan hasil tes fisik seorang atlet beladiri. Berikut Gambaran tahapan-tahapan yang dapat dilihat pada Gambar 1.



Gambar 1 Tahapan proses penelitian

**2.1 Indentifikasi Data**

Data yang ada adalah digunakan untuk mengklasifikasi nilai data fisik seorang atlet beladiri.

**2.2 Pengumpulan Data**

Data yang digunakan adalah data yang

didapat dari pelatih fisik yang telah di tes atlet saat di KONI.

Tabel 1 Data yang digunakan

P4	P3	P2	P1	T4	T3	T2	T1	Total
1	3	3	3	4	9	6	3	22
2	4	4	0	8	12	8	0	28
0	2	7	1	0	6	14	1	21
3	4	3	0	12	12	6	0	30
1	1	5	3	4	3	10	3	20
4	0	5	1	16	0	10	1	27
2	1	3	4	8	3	6	4	21
7	3	0	0	28	9	0	0	37
3	3	2	2	12	9	4	2	27
5	3	2	2	20	9	4	2	35
2	4	4	0	8	12	8	0	28
3	2	4	1	12	6	8	1	27
5	3	1	1	20	9	2	1	32
3	1	6	0	12	3	12	0	27
2	1	4	3	8	3	8	3	22
3	2	5	0	12	6	10	0	28
2	5	2	1	8	15	4	1	28
2	3	4	1	8	9	8	1	26
0	1	6	3	0	3	12	3	18
2	0	4	4	8	0	8	4	20
4	3	2	1	16	9	4	1	30
4	2	2	2	16	6	4	2	28
6	2	2	0	24	6	4	0	34
0	2	4	4	0	6	8	4	18
1	2	3	4	4	6	6	4	20
1	2	4	3	4	6	8	3	21
2	1	4	3	8	3	8	3	22
5	2	1	2	20	6	2	2	30
5	2	3	0	20	6	6	0	32
1	2	5	2	4	6	10	2	22
3	1	3	3	12	3	6	3	24
2	2	4	2	8	6	8	2	24
2	0	4	4	8	0	8	4	20
2	2	2	4	8	6	4	4	22
1	1	3	5	4	3	6	5	18
3	2	5	0	12	6	10	0	28
1	3	3	3	4	9	6	3	22
4	1	4	1	16	3	8	1	28
2	3	3	2	8	9	6	2	25
3	3	3	1	12	9	6	1	28
7	1	2	0	28	3	4	0	35
4	1	3	2	16	3	6	2	27
2	1	3	4	8	3	6	4	21
3	1	3	3	12	3	6	3	24
2	4	2	2	8	12	4	2	26
2	2	3	3	8	6	6	3	23

Berikut keterangan dari variable Tabel 1 :

- P4 = Point maksimal
- P3 = Point kurang maksimal
- P2 = Point Minimal
- P1 = Point tidak tercapai
- T4 = Total point (x4)
- T3 = Total point (x3)
- T2 = Total point (x2)
- T1 = Total point (x1)
- Total = Hasil point yang sudah dikali

### 2.3 Perhitungan Naïve Bayes

Perhitungan data menggunakan metode Naïve Bayes Langkah selanjutnya adalah pengolahan data menggunakan metode Naïve Bayes. Langkah ini melakukan perhitungan dengan menggunakan rumus *Naïve Bayes* dari data yang telah didapat sebelumnya[6].

### 2.4 Penggunaan Aplikasi

Lalu selanjutnya adalah penggunaan aplikasi Weka. Data yang didapat dimasukan kedalam aplikasi Weka dan di proses dengan operasi Naïve Bayes.

### 2.5 Kesimpulan

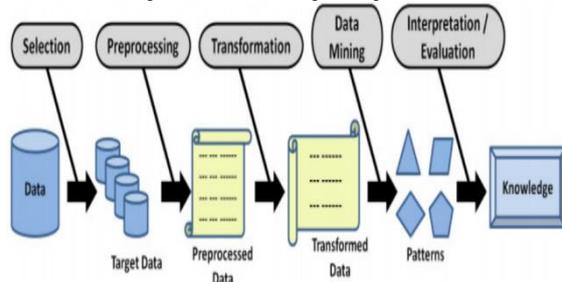
Kesimpulan dijelaskan setelah didapatkan hasil dari pengolahan data aplikasi penggunaan aplikasi.

### 2.6 Data Mining

*Data mining* adalah aktivitas yang mengambarkan sebuah proses analisis yang terjadi secara iteratif pada *database* yang besar, dengan tujuan mengekstrak informasi dan *knowledge* yang akurat dan berpotensi berguna untuk *knowledge workers* yang berhubungan dengan pengambilan keputusan dan pemecahan masalah. *Data mining* merupakan salah satu solusi untuk menjelaskan proses penggalian informasi dalam suatu basis data yang berskala besar dan proses klasifikasi otomatisasi kasus berdasarkan pola data yang diperoleh dari data set[6].

### 2.7 Knowledge Discovery In Database (KDD)

KDD adalah metode yang digunakan untuk dapat memperoleh pengetahuan yang berasal dari database yang ada. Hasil pengetahuan yang diperoleh dapat dimanfaatkan untuk basis pengetahuan (*knowledge base*) yang digunakan dalam keperluan mengambil keputusan[7]. Secara lebih detail, proses KDD seperti pada Gambar 2.1.



Gambar 2 Proses Knowledge Discovery In Database (KDD)

Penjelasan dari Gambar 2 dilakukan dengan beberapa langkah

#### a) Selection

Selection digunakan untuk menentukan variabel yang akan diambil agar tidak ada kesamaan dan

terjadi perulangan yang tidak diperlukan dalam pengolahan data mining[8].

Tabel 2 Pra-proses data jenis kelamin

jenis kelamin	
laki-laki	30
perempuan	16

Dalam Tabel 2 data jenis kelamin laki-laki berjumlah 30 orang dan kelamin perempuan berjumlah 16.

Tabel 3 Pra-proses data berat badan

berat badan	
laki-laki	55-70 kg
perempuan	45-60 kg

Tabel 3 menjelaskan berat badan laki-laki mulai dari berat minimal 55kg hingga maksimal 70kg, perempuan mulai berat minimal 45kg hingga maksimal 60kg.

Tabel 4 data test fisik

vertical jump	
laki-laki	min 50 cm
perempuan	min 40 cm

Tabel 4 menjelaskan test fisik vertical jump laki-laki minimal 50cm dan maksimal tidak ada batasnya, sedangkan perempuan minimal 40cm dan maksimal tidak ada batasan.

Table 5 data test fisik

20m sprint	
laki-laki	3,2 detik
perempuan	4 detik

Table 6 data test fisik

sit up	
laki-laki	40 kali
perempuan	35 kali

Tabel 6 data test fisik sit up laki-laki minimal 40 kali dalam 1 menit, sedangkan perempuan 35 kali dalam 1 menit[9].

Tabel 7 data test fisik

push up	
laki-laki	40 kali
perempuan	35 kali

Tabel 7 data test fisik push up dilakukan dalam 1 menit yang berjumlah laki-laki minmal 40kali dalam 1 menit, dan perempuan 35 kali dalam 1 menit.

Table 8 data status

status	
mahasiswa	31
pekerja	15

b) *Preprocessing*

Pada preprocessing terdapat dua tahap, yaitu sebagai berikut :

1. *Data Cleaning*

Menghilangkan data yang tidak diperlukan seperti menangani missing value, noise data serta menangani data-data yang tidak konsisten dan relevan.

2. *Data Intergation*

Dilakukan terhadap atribut yang mengidentifikasi entitas yang unik

c) *Transformation*

Mengubah data sesuai format ekstention yang sesuai dalam pengolahan data mining karena beberapa metode pada data mining memerlukan format khusus sebelum dapat diproses pada data mining.

d) *Data Mining*

Proses utama pada metode yang diterapkan untuk mendapatkan pengetahuan baru dari data yang diproses[10]. Pada penelitian ini diterapkan teknik klasifikasi yaitu metode *Naïve Bayes*.

e) *Evaluation/Interpretation*

Mengidentifikasi pola-pola yang menarik kedalam knowdge base yang diidentifikasi. Pada tahap ini, menghasilkan pola-pola khas maupun model prediksi yang dievaluasi untuk menilai kajian yang ada sudah memenuhi target yang diinginkan.

f) *Knowledge*

Pola-pola yang dihasilkan akan dipresentasikan kepada pengguna. Pada tahapan ini pengetahuan baru yang dihasilkan bisa dipahami semua orang yang akan dijadikan acuan pengambilan keputusan[11].

**2.8 Klasifikasi**

Proses penemuan model (atau fungsi) yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui. Algoritma klasifikasi yang banyak digunakan secara luas, yaitu *Decision/classification trees, Bayesian classifiers/ Naïve Bayes classifiers, Neural networks, Analisa Statistik, Algoritma Genetika, Rough sets, k-nearest neighbor, Metode Rule Based, Memory based reasoning, dan Support vector machines (SVM)* [6][12].

**2.9 Kondisi Fisik**

Dalam teori latihan, disebutkan ada empat

aspek latihan yang perlu diperhatikan dan dilatih secara seksama yaitu latihan fisik, latihan teknik, latihan taktik, dan latihan mental. Kondisi fisik merupakan salah satu aspek penting untuk meningkatkan prestasi olahraga, dengan melakukan latihan kondisi fisik memungkinkan atlet untuk dapat mengikuti program latihan dengan baik. Kondisi fisik merupakan salah satu aspek latihan yang paling dasar untuk dilatih dan di tingkatkan, untuk mendapatkan kondisi fisik yang baik diperlukan persiapan latihan yang dapat meningkatkan dan mengembangkan kondisi fisik, daya tahan merupakan salah satu komponen fisik yang sangat penting untuk dilatih dan ditingkatkan menjadi stamina dalam upaya mencapai prestasi yang optimal [4].

**2.10 Algoritma Naïve Bayes**

Bayes merupakan pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu class. Bayes memiliki akurasi dan kecepatan yang sangat tinggi saat diaplikasi ke dalam *database* dengan data yang besar. Formula Naïve Bayes pada Persamaan 1 [6][13] :

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

Notasi persamaan 1 dapat dijelaskan:

X = Data dengan class yang belum diketahui

H = Hipotesis data X merupakan suatu class spesifik

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi x (*posteriori probability*)

P(H) = Probabilitas hipotesis H (*prior probability*)

P(X|H) = Probabilitas X berdasarkan kondisi tersebut

P(X) = Probabilitas dari X

**2.11 Sistem Pendukung Keputusan**

Sistem itu digunakan untuk membantu pengambilan keputusan dalam situasi yang semi terstruktur dan situasi yang tidak terstruktur. Dikatakan bahwa aplikasi sistem pendukung keputusan haruslah fleksibel, interaktif, adaptif, dan memiliki antarmuka yang mudah (user friendly). Sistem pendukung keputusan tidak dimaksudkan untuk mengotomatisasikan pengambilan keputusan, tetapi memberikan perangkat interaktif yang memungkinkan pengambil keputusan untuk melakukan berbagai analisis[14][15][16].

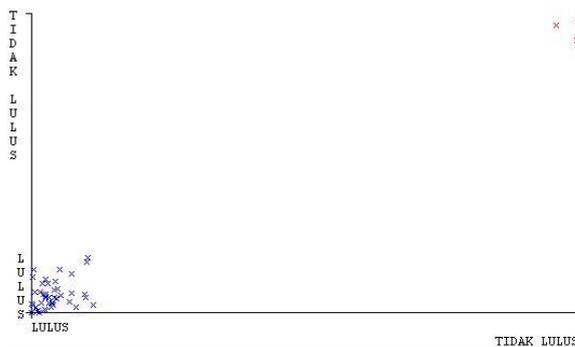
**3. Hasil Dan Pembahasan**

Pada penelitian ini proses klasifikasi *data mining Metode Naïve Bayes* menggunakan aplikasi weka, aplikasi weka berfungsi untuk membantu proses klasifikasi data dengan tool Classify dan menggunakan salah satu metode algoritma Naïve Bayes. Sampel data yang akan digunakan dalam penelitian ini sebanyak 46 data yang sudah ditata sesuai atribut yang diperlukan untuk proses data mining, setelah itu akan dikelompokkan menjadi 2

classify yaitu Data Lulus dan tidak lulus test sesuai atribut yang telah didapat sebelumnya. Berikut hasil dari proses *data mining*.

Tabel 7 Hasil Use Data Training Set

Klasifikasi	Jumlah	Persen
Lulus	43	93.5%
Tidak Lulus	3	6.5%



Gambar 3 Grafik Visual Hasil Classify

Pada proses uji coba ini dilakukan dengan menggunakan seluruh data. Dari proses sebanyak 46 data diperoleh hasil nilai classify.

Tabel 8 Hasil Percentage Split 60%

Klasifikasi	Jumlah	Persen
Lulus	17	94.4%
Tidak Lulus	1	5.6%

Tabel 8 adalah hasil dari proses percentage split 60% Correctly Classified Instances 94,4%, Incorrectly Classified Instances 5,6%, Precision 0.944, Recall 1.000. Perbedaan antara percentage split 70% pada Tabel 9.

Tabel 9 Hasil percentage split 70%

Klasifikasi	Jumlah	Persen
Lulus	13	92.9%
Tidak Lulus	1	7.1%

Pada Tabel 9 hasil dari proses percentage split 60% Correctly Classified Instances 92,9%, Incorrectly Classified Instances 7,1%, Precision 0.929, Recall 1.000.

Tabel 10 Hasil percentage split 80%

Klasifikasi	Jumlah	Persen
Lulus	8	88.9%
Tidak Lulus	1	11.1%

Pada Tabel 10 hasil dari proses percentage split 60% Correctly Classified Instances 88,9%, Incorrectly Classified Instances 11,1%, Precision 0.889, Recall 1.000. Dari hasil perhitungan Klasifikasi data traning tersebut menghasilkan perbedaan hasil dari Tabel ke 7 hingga Tabel 10. Hasil Tabel ke-7 berjumlah 43 yang lulus, Tabel ke-8 berjumlah 18 yang lulus dengan menggunakan percentage split 60%, Tabel 9 berjumlah 13 yang lulus menggunakan percentage split 70%, sedangkan Tabel ke-10 berjumlah 8 yang lulus menggunakan percentage split 80%.

Dapat dilihat pada Gambar 3 hasil dari perluasan data klasifikasi berdasarkan tingkat kelulusan yang telah dihitung. Pada Gambar 3 memperlihatkan bahwa klasifikasi kelulusan berada diposisi bawah merupakan kelompok data yang memiliki nilai Kelulusan hasil test fisik, sedangkan klasifikasi tidak lulus berada di posisi atas merupakan kelompok data yang memiliki nilai tidak lulus.

#### 4. Kesimpulan

Dari penelitian ini terbukti bahwa model algoritma naïve bayes mempunyai hasil yang baik, dari hasil evaluasi data yang digunakan mampu menganalisis tingkat kelulusan atlet beladiri kempo dalam melakukan test fisik. Dari data test fisik beladiri kempo dapat dinyatakan data set yang mempunyai akurasi sangat baik. Banyak atribut pada data set tersebut bisa dikelola dengan baik dan mempengaruhi tingkat akurasi dari model algoritmat Naïve Bayes.

#### Daftar Pustaka

- [1] E. K. br Meliala, "Analisis Kondisi Fisik Atlet Putra Floorball Universitas Negeri Surabaya," *JOSSAE J. Sport Sci. Educ.*, vol. 3, no. 2, p. 81, 2019, doi: 10.26740/jossae.v3n2.p81-93.
- [2] C. F. Villareal, J. U. Rojas, and M. X. Ochoa, "Impacto operacional de los equipos caninos en la lucha contra el narcotráfico en Colombia: una revisión de la contribución, estrategias y tendencias recientes," *Rev. Sci. Tech.*, vol. 37, no. 1, pp. 189-200, 2018, doi: 10.20506/rst.37.1.2751.
- [3] C. W. Kuswanto, "Penyusunan tes fisik atlet pencak silat dewasa kategori tanding," *J. Keolahragaan*, vol. 4, no. 2, p. 145, 2016, doi: 10.21831/jk.v4i2.6423.
- [4] T. D. Putranto, F. Y. Wulandari, and A. Sifaq, "Profil kondisi fisik atlet panahan gresik," *J. Sport Exerc. Sci.*, vol. 1, no. 1, pp. 20-24, 2018.
- [5] E. Science, "PENGARUH LATIHAN BOUNCE BALL 2 VS 2 DAN 3 VS 3 TERHADAP

- KELINCAHAN,” vol. 1, no. 2, pp. 1–5, 2018.
- [6] H. Annur, “Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes,” *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.
- [7] M. H. Rifqo and A. Wijaya, “Implementasi Algoritma Naive Bayes Dalam Penentuan Pemberian Kredit,” *Pseudocode*, vol. 4, no. 2, pp. 120–128, 2017, doi: 10.33369/pseudocode.4.2.120-128.
- [8] A. Saleh, “Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga,” *Creat. Inf. Technol. J.*, vol. 2, no. 3, pp. 207–217, 2015.
- [9] D. P. Saputro and S. Siswantoyo, “Norm drafting of pencak silat physical test for adolescent fighting category,” vol. 6, no. 1, pp. 1–10, 2018.
- [10] A. Nugroho, R. Hidayatillah, S. Sumpeno, and M. H. Purnomo, “Klasifikasi Interaksi Kampanye di Media Sosial Menggunakan Naive Bayes Kernel Estimator,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 8, no. 2, p. 107, 2019, doi: 10.22146/jnteti.v8i2.499.
- [11] D. Evanko, “Optical imaging of the native brain,” *J. Article*, vol. 7, no. 1, p. 34, 2010, doi: 10.1038/nmeth.f.284.
- [12] B. Wijaya and A. N. Tresna Maulana, Fahrudin, “Analisis Dan Pemetaan Jumlah Penumpang Kereta Api Di Indonesia Menggunakan Metode Statistik Deskriptif Dan K-Means Clustering,” *J. Mantik*, vol. 3, no. 2, 2019, Accessed: Jun. 22, 2020. [Online]. Available: <http://iocscience.org/ejournal/index.php/mantik/article/view/236>.
- [13] M. Halim, “Pembuatan Aplikasi Pendukung Keputusan Beasiswa Tidak Mampu Dengan Metode Naive Bayes,” *JSAI (Journal Sci. Appl. Informatics)*, vol. 1, no. 1, pp. 24–31, 2018, doi: 10.36085/jsai.v1i1.6.
- [14] P. Soepomo, “Perancangan Sistem Pendukung Keputusan Penentuan Rumah Sehat,” *Peranc. Sist. Pendukung Keputusan Penentuan Rumah Sehat*, vol. 1, no. 2, pp. 584–596, 2013, doi: 10.12928/jstie.v1i2.2582.
- [15] A. . Fallis, “濟無NFallis, A. . (2013) 濟無No Title No Title’, *Journal of Chemical Information and Modeling*, 53(9), pp. 1689–1699. doi: 10.1017/CBO9781107415324.004.o Title No Title,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013, doi: 10.1017/CBO9781107415324.004.
- [16] H. Rohayani, “Analisis Sistem Pendukung Keputusan Dalam Memilih Program Studi Menggunakan Metode Logika Fuzzy,” *J. Sist. Inf.*, vol. 5, no. Analisis Sistem Pendukung Keputusan, pp. 530–539, 2013.

# SYSTEMIC: Information System and Informatics Journal

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 1 - Agustus 2020

## Sistem Penunjang Keputusan Pemasaran Produk X Menggunakan Metode K-Means

Ach. Syuhbanul Yaumi<sup>1</sup>, Zainul Zulfikar<sup>2</sup>, Aryo Nugroho<sup>3</sup><sup>1,2,3</sup> Universitas Narotama, Surabaya[ach.syuhbanul@mhs.fasilkom.narotama.ac.id](mailto:ach.syuhbanul@mhs.fasilkom.narotama.ac.id)<sup>1</sup>, [zainul.16@fik.narotama.ac.id](mailto:zainul.16@fik.narotama.ac.id)<sup>2</sup>, [aryo.nugroho@narotama.ac.id](mailto:aryo.nugroho@narotama.ac.id)<sup>3</sup>

### Kata Kunci

Data mining,  
Klasterisasi, K-Means,  
Pemasaran, Produk X.

### Abstrak

Permasalahan yang sedang ada di toko saat ini adalah kesulitan untuk mengetahui produk X yang saat ini banyak diminati atau paling banyak digunakan oleh konsumen agar persediaan produk x dapat dipenuhi sesuai permintaan pelanggan dan tidak terjadi out of stock. Oleh karena itu, dalam penelitian pengelompokan dengan metode K-Means untuk pemasaran produk x merupakan salah satu cara untuk mengetahui pilihan pelanggan terhadap produk x yang dikonsumsi. Pada penelitian ini melakukan pengelompokan dari data hasil penyebaran angket atau kusioner yang disebar di toko, kemudian data tersebut di kelompokkan menjadi 2 kelompok menggunakan salah satu algoritma klasterisasi yaitu K-Means. Data yang digunakan merupakan data yang dikumpulkan sebanyak 366 data responden pelanggan toko. Setelah data tersebut diproses menggunakan salah satu metode data mining yaitu algoritma K-Means, menunjukkan bahwa cluster 1 merupakan kelompok konsumen tipe A dengan presentase sebesar 33%, sedangkan cluster 2 merupakan kelompok konsumen tipe B dengan presentase sebesar 67%.

### Keywords

Clustering, Data mining,  
Marketing, K-Means,  
Products X.

### Abstract

The problem that is currently in store is the difficulty to find out which product x is currently in high demand or is most widely used by consumers so that inventory of product x can be met according to customer demand and does not occur out of stock. Therefore, in the research grouping with the K-Means method for marketing product x is one way to determine customer choices for product x consumed. In this study grouping data from questionnaires or questionnaires that are distributed in stores, then the data are grouped into 2 groups using one of the clustering algorithms, K-Means. The data used are data collected by 366 respondents of store customers. After the data is processed using one of the data mining methods, the K-Means algorithm, shows that cluster 1 is a type A consumer group with a percentage of 33%, while cluster 2 is a type B consumer group with a percentage of 67%.

## 1. Pendahuluan

Pada umumnya produk x tipe A biasanya terbuat dari bahan tembakau yang digulung oleh kertas. Namun seiring berkembangnya zaman terdapat jenis produk x baru yaitu tipe B. Banyak yang mengira bahwa produk x tipe B lebih aman dibandingkan dengan produk x tipe A. Beberapa alasan dari mereka yang beralih dari tipe A ke ke tipe B juga bermacam-macam, mulai dari aman bagi kesehatan, lebih ramah lingkungan karena menghasilkan uap air bukan asap, mengurangi resiko dari bahaya penyakit, mengurangi kecanduan, dan yang paling utama sebagai alternative pengganti produk tipe A untuk perlahan-lahan berhenti kecanduan.

Beberapa diantara pengkonsumsi produk tipe B adalah usia remaja yang pada dasarnya merupakan masa-masa ingin mencari tahu dan mencoba hal-hal baru tanpa berpikir rasional. Dari hasil penyebaran kusioner di dapat berbagai rentang usia yang telah dibagi menjadi beberapa status dengan jumlah 49,5% dari kalangan mahasiswa, 39% dari kalangan pekerja, dan 11,5% dari kalangan pelajar.

Di era globalisasi, teknologi semakin berkembang bahkan perkembangan teknologi saat ini banyak membantu memudahkan pekerjaan manusia supaya lebih efisien dan efektif dalam berbagai bidang [1]. Klasterisasi adalah salah satu metode *data mining* yang sangat membantu dalam

penemuan pengetahuan [2]. Klusterisasi berarti mengelompokkan objek data ke dalam sebuah kelompok yang memiliki paling sedikit kesamaan dengan anggota *cluster* lainnya [3].

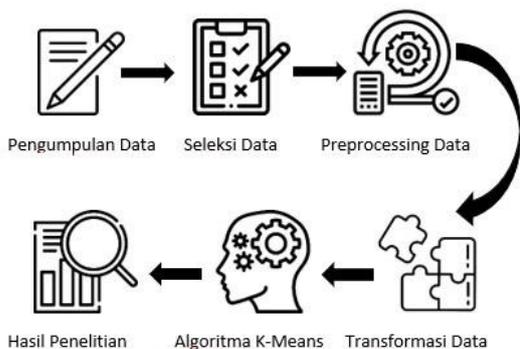
Permasalahan yang sedang ada di toko saat ini adalah kesulitan mengetahui produk x yang saat ini banyak diminati atau paling banyak digunakan oleh konsumen agar persediaan produk x dapat dipenuhi sesuai permintaan pelanggan dan tidak terjadi *out of stock*. Toko ini membutuhkan suatu sistem pendukung keputusan informasi penting yang nantinya bisa bermanfaat untuk penjualan agar dapat memenuhi persediaan produk x yang sesuai permintaan pelanggan. Metode ini digunakan karena mampu menghasilkan kriteria yang telah ditetapkan berdasarkan seleksi keputusan terbaik dari sejumlah keputusan [4].

Sistem Pendukung Keputusan merupakan suatu sistem informasi interaktif yang menyediakan pemodelan, manipulasi data dan informasi [5]. Dalam proses pengambilan keputusan pengetahuan dan pengalaman merupakan peran penting [6].

Oleh karena itu dalam penelitian Sistem Pendukung Keputusan pengelompokkan dengan *K-Means* untuk pemasaran produk x merupakan salah satu cara untuk mengetahui selera atau pilihan pelanggan di toko ini terhadap produk x yang dikonsumsi. Karena metode *K-Means* mempunyai hasil pembagian yang lebih akurat [5]. Keakuratan metode K-Means tergantung pada pemilihan data centroid selama inisialisasi [7]. Tujuan dari analisis pengelompokkan pemasaran produk x adalah untuk merancang strategi dalam penjualan produk x di toko ini pada masa yang akan mendatang berdasarkan dua *cluster* yaitu *cluster 1* adalah produk x tipe A sedangkan *cluster 2* adalah produk x tipe B.

## 2. Metode Penelitian

Dalam sebuah penelitian, metode berfungsi sebagai cara untuk memecahkan masalah dan menggambarkan tahapan-tahapan yang akan dilakukan. Pada penelitian ini, untuk menganalisis dan menentukan tipe produk x yang digunakan. Berikut Gambaran tahapan-tahapan yang dilakukan dapat dilihat pada Gambar 1.



Gambar 1 Tahapan Penelitian

### 2.1 Pengumpulan Data

Pengumpulan data berarti mengumpulkan data yang dibutuhkan untuk analisis[8]. Pada

tahap ini bertujuan untuk mencari atau memperoleh informasi bagi pengguna produk x sebagai bahan penelitian yang akan di klusterisasi dan dihitung menggunakan metode *K-means* yang di dapat dari penyebaran angket kusioner melalui *google form* atau *hard copy* kepada pembeli atau hanya sekedar bersantai di toko.

### 2.2 Seleksi Data

Fungsi dari fitur seleksi digunakan untuk meringankan beban pemrosesan dalam *data mining*[9]. Tujuannya agar mendapatkan hasil *data mining* yang spesifik dan akurat, hanya beberapa data yang diambil dari data mentah dan akan digunakan sebagai atribut untuk selanjutnya diproses. Ada 7 atribut yang digunakan yaitu status, usia, mulai mengkonsumsi, merk produk, alasan memilih produk x, pemakain perhari, dan pengeluaran perbulan. Data yang dikumpulkan sebanyak 366 data responden yang merupakan pelanggan di toko.

### 2.3 Preprocessing Data

Data mentah yang telah didapat dari hasil penyebaran angket atau kusioner masih perlu diproses lagi untuk memperbaiki data yang belum sempurna dan data yang masih kurang atau belum lengkap. Tujuan dari *preprocessing* data yaitu pembersihan atau pengurangan atribut data yang tidak diperlukan serta kolom variabel yang tidak memiliki nilai, selain itu *preprocessing* data juga meningkatkan kualitas data dan mengurangi ukuran file log web[10]. Data yang kurang lengkap dan tidak relevan hanya akan membuat proses perhitungan pada weka mengalami *error*, oleh karena itu diperlukan *preprocessing* data yang akan digunakan untuk kebutuhan *data mining*.

Tabel 1 Data yang digunakan

No	JK	S	U	MM	MP	AM	PP	PB
1	1	3	22	17	2	1	1	1
2	1	1	17	13	3	2	1	1
3	1	3	21	19	4	3	1	1
4	1	2	21	19	4	3	1	1
5	1	2	19	15	4	1	2	2
6	1	3	22	16	1	1	2	2
7	1	2	20	15	5	4	1	1
8	1	1	17	16	3	2	1	1
9	1	3	24	19	1	4	2	1
10	1	3	21	13	10	3	1	1
11	1	2	26	20	1	2	3	1
12	1	1	17	12	1	2	1	1
13	1	1	18	16	6	4	1	1
14	1	2	18	15	1	3	1	1

Berikut keterangan dari variable Tabel 1:

- JK = Jenis Kelamin
- S = Status
- U = Usia
- MM = Mulai Mengonsumsi
- MP = Merk Produk
- AM = Alasan Memilih
- PP = Pengeluaran Perhari
- PB = Pengeluaran Bulanan

## 2.4 Transformasi Data

Proses ini bertujuan untuk mengubah data yang dapat diolah dengan algoritma yang akan digunakan dalam bentuk klasifikasi atau numerik[11]. Data yang diperoleh dari hasil penyebaran angket atau kusioner belum sesuai format untuk digunakan pemrosesan data. Pemrosesan data diperlukan *input* data yang sudah dikategorikan dan tipe data *numerik*, maka jenis data nominal seperti status, usia, mulai mengonsumsi, merk produk x, alasan memilih produk x, pemakain perhari, dan pengeluaran perbulan terlebih dahulu diinisialisasikan dalam bentuk angka.

Tabel 2 Pra-Proses data Status

Tipe Status	Transformasi
Pelajar	1
Mahasiswa	2
Pekerja	3

Tabel 3 Pra-Proses data Merk Produk

Merk Produk	Transformasi
Surya, Gudang Garam	1
U Mild	2
Mild	3
Malboro	4
Dunhill	5
Djarum	6
Sukun	7
Marshband	8
Classmild	9
Sampoerna,Samsu	10
Pro Mild	11
Tesla	12
Pod,Smoke	13
Mod,Hxhom	14
Vgod	15
Aegis	16
Invander 3	17
Mekanika	18
Geekvape	19
Lswtch	20
Uwel Caliburn	21
Battlestar	22

Tabel 4 Pra-Proses data Alasan Memilih

Alasan	Transformasi
Lebih Enak,Enak,Suka	1
Hemat,Lebih Hemat	2
Lebih Sempel,Praktis	3
Lebih Nikmat, Lebih Berasa	4
Lebih Aman,Lebih Sehat	5
Faktor Lingkungan	6
Karena Rasa,Banyak Rasa	7
Mengurangi Merokok	8
Gaya,Kekinian,Coba-Coba	9

Tabel 5 Pra-Proses data Pemakain Perhari

Pemakain Perhari	Transformasi
1-10 Batang	1
10-20 Batang	2
20-30 Batang	3
1-10 ml	4
10-20 ml	5
20-30 ml	6

Tabel 6 Pra-Proses Pengeluaran Perbulan

Kisaran Nominal	Transformasi
<500 ribu	1
>500 ribu	2

## 2.5 Perhitungan data dengan *K-Means*

Algoritma *K-Means* merupakan salah satu metode *clustering* yang paling populer[12]. Tujuan dari algoritma ini adalah untuk membagi data menjadi beberapa kelompok[13]. Algoritma *K-Means* ini akan menerima masukan berupa data tanpa label kelas. Pada penelitian ini perhitungan *K-Means* menggunakan aplikasi weka, Data yang dipilih dari data mentah melalui proses seleksi data, *processing* data dan *transformasi* data akan dimasukkan lalu membaca data *training* dan diproses, kemudian menilai hasil dari proses tersebut. Untuk menghitung iterasi jarak terdekat antar *cluster* menggunakan Persamaan (1) *Euclidean Distance*. Dengan menentukan nilai *centroid* awal (nilai tengah) kemudian partisi data menggunakan nilai *minimum*. Lakukan iterasi sampai nilai atau posisi data tidak mengalami perubahan[14][15]. Jika data sudah tidak berubah lagi maka hentikan iterasi. Maka hasil dari perhitungan iterasi merupakan *centroid* akhir dari data yang sudah dipartisi sebelumnya. Berikut adalah *flowchart* diagram algoritma *K-Means* :



Gambar 2 Flowchart Algoritma K-Means

Algoritma K-Means Klasterisasi dapat dijelaskan sebagai berikut :

1. Tentukan *k* buah *cluster*
2. Pilih sejumlah *k* buah objek secara acak yang akan dijadikan sebagai titik centroid *cluster*
3. Tentukan *k* buah centroid (titik tengah)
4. Kelompokkan objek ke centroid *cluster* terdekat berdasarkan Euclidean Distance. Berikut dibawah ini rumus *Euclidean Distance*:
- 5.

$$D(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots \dots \dots (1)$$

Notasi Persamaan 1 dapat dijelaskan sebagai berikut :

- D = jarak antara x dan y
- x = data pada atribut
- y = nilai pusat
- j = setiap data
- n = jumlah data
- x<sub>j</sub> = nilai pusat ke j
- y<sub>j</sub> = objek data ke j

6. Hitung kembali semua nilai titik centroid Ulangi langkah 3 sampai 5 hingga nilai titik centroid tidak lagi berubah

### 2.6 Hasil Penelitian

Pada tahap terakhir ini bertujuan untuk

memvisualisasikan hasil dari proses *data mining* berupa Tabel atau Gambar serta menganalisa dan mengevaluasi apakah dapat digunakan untuk menyelesaikan masalah pada penelitian dengan hasil *data mining*.

### 3. Hasil Dan Pembahasan

Pada penelitian ini proses *data mining* menggunakan aplikasi weka , aplikasi tersebut berfungsi untuk membantu proses pengelompokan data dengan tool cluster dan menggunakan salah satu metode algoritma *K-Means* dari *cluster* itu sendiri. Sampel data yang akan digunakan dalam penelitian ini sebanyak 366 data yang sudah ditata sesuai atribut yang diperlukan untuk proses data mining, setelah itu akan dikelompokkan menjadi 2 *cluster* yaitu pengguna produk x tipe A atau produk x tipe B sesuai atribut-atribut yang telah didapat sebelumnya. Berikut hasil dari proses *data mining*.

Tabel 7 Hasil Use Data Training Set

Attribute	Full Data	Cluster 0	Cluster 1
	(366)	(122)	(244)
Status	2.224	2.2213	2.2254
Usia	22.4863	22.5246	22.4672
Mulai Mengkonsumsi	17.3497	15.7131	18.168
Merk Produk	11.9481	4.5	15.6721
Alasan Memilih	4.2814	2.8361	5.0041
Pemakaian Perhari	3.3169	1.1885	4.3811
Pengeluaran Bulanan	1.5055	1.6311	1.4426

Pada proses pengetesan ini dilakukan dengan menggunakan seluruh data *training* itu sendiri. Dari proses sebanyak 366 data diperoleh hasil nilai *cluster* 0 sebanyak 33% dengan jumlah data sebesar 122 data dan *cluster* 1 sebanyak 67% dengan jumlah data sebesar 244 data.

Tabel 8 Hasil Percentage Split 85%

Attribute	Full Data	Cluster 0	Cluster 1
	(311)	(159)	(152)
Status	2.2669	2.4088	2.1184
Usia	22.6785	23.956	21.3421
Mulai konsumsi	17.4695	17.2264	17.7237
Merk Produk	12.0579	11.2453	12.9079
Alasan Memilih	4.2926	4.3711	4.2105
Pemakaian Perhari	3.3312	3.1887	3.4803
Pengeluaran Bulanan	1.5113	2	1

Tabel 8 adalah hasil dari proses percentage split dimana hasil klasterisasi dipes menggunakan nilai k% dari data yang telah diinput, nilai k adalah nilai masukan yang ditentukan oleh user. Dalam proses uji tes ini menggunakan nilai k sebesar 85%. Artinya nilai 85% datanya akan menjadi *training set* sedangkan sisanya yaitu 15% menjadi *test set*. Dan didapatkan hasil *cluster 0* sebanyak 47% dan *cluster 1* sebanyak 53%.

Sedangkan proses pengelompokan data secara manual dengan rumus *Euclidean Distance* diawali dengan menentukan nilai *cluster*. Jumlah *cluster* yang akan digunakan pada penelitian ini adalah 2 buah *cluster*, sehingga nilai  $k=2$  yaitu *cluster 1* dan *cluster 2*. Kemudian tentukan nilai titik tengah *cluster* (centroid). Pada tahap ini nilai centroid dapat dipilih secara acak dari *dataset* yang akan digunakan. Hasil penentuan nilai centroid awal dapat dilihat pada Tabel 9.

Tabel 9 Centroid Awal

Cluster 1	Cluster 2
1	1
2	2
21	19
19	18
4	14
3	7
1	4
1	1

Pada Tabel 9 Nilai centroid awal ditentukan secara acak, diambil dari data mentah yang sudah diubah menjadi data *numerik*. Untuk proses selanjutnya dilakukan pengulangan ke- 1 sampai selesai. Jika centroid baru tidak sama dengan centroid sebelumnya, maka proses pengulangan tetap dilanjutkan, tetapi jika centroid baru yang dihitung sama dengan centroid sebelumnya, maka proses klasterisasi telah selesai.

Tabel 10 Hasil Iterasi ke-1

data ke -I	c1	c2	Min	cluster
1	3.74	14.14	3.74	1
2	7.42	13.60	7.42	1
3	1.00	11.45	1.00	1
4	0.00	11.40	0.00	1
5	5.10	12.25	5.10	1
6	5.10	14.97	5.10	1
7	4.36	10.44	4.36	1
8	5.29	12.81	5.29	1
9	4.58	14.46	4.58	1
10	8.54	8.43	8.43	2
11	10.86	5.29	5.29	2
12	16.70	11.79	11.79	2
13	10.95	3.46	3.46	2
14	16.25	5.48	5.48	2
15	15.72	14.46	14.46	2

Tabel 11 Hasil Iterasi ke-2

data ke -I	c1	c2	Min	cluster
1	2.55	13.57	2.55	1
2	5.05	15.00	5.05	1
3	2.94	11.35	2.94	1
4	2.84	11.35	2.84	1
5	2.84	12.73	2.84	1
6	3.25	14.48	3.25	1
7	3.09	11.22	3.09	1
8	3.63	14.15	3.63	1
9	5.23	13.69	5.23	1
10	7.96	8.41	7.96	1
11	11.86	4.59	4.59	2
12	18.09	6.98	6.98	2
13	11.93	4.33	4.33	2
14	16.96	7.33	7.33	2
15	17.99	10.74	10.74	2

Tabel 12 Hasil Iterasi ke-3

data ke-I	c1	c2	min	Cluster
1	2.94	14.86	2.94	1
2	4.93	16.53	4.93	1
3	3.04	12.53	3.04	1
4	2.98	12.53	2.98	1
5	2.58	14.19	2.58	1
6	3.67	15.79	3.67	1
7	2.46	12.71	2.46	1
8	3.75	15.53	3.75	1
9	5.63	14.71	5.63	1
10	7.16	10.09	7.16	1
11	11.32	5.30	5.30	2
12	17.55	6.06	6.06	2
13	11.33	5.22	5.22	2
14	16.29	7.57	7.57	2
15	17.78	9.71	9.71	2

Setelah dilakukan perhitungan iterasi sebanyak 4 kali dapat dilihat pada Tabel 10,11,12 dan 13. Dari perhitungan menggunakan rumus *Euclidean Distance* didapatkan hasil nilai centroid akhir pada iterasi ke-4 dimana nilai centroid tidak berubah dari hasil perhitungan sebelumnya. Di dapatkan hasil nilai centroid akhir dari perhitungan iterasi dapat dilihat sebagai berikut.

Tabel 13 Hasil Iterasi ke-4

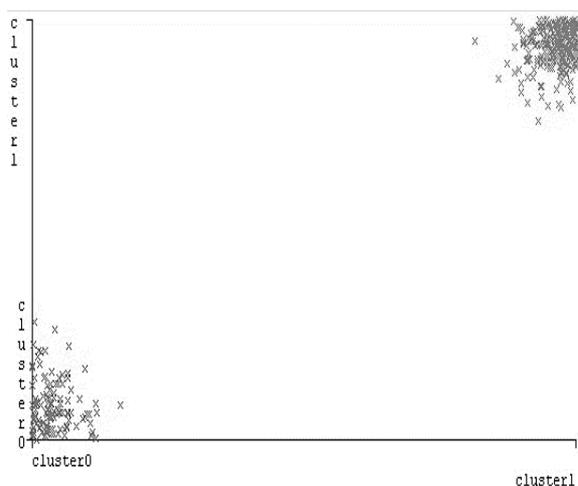
data ke-I	c1	c2	min	cluster
1	2.94	14.86	2.94	1
2	4.93	16.53	4.93	1
3	3.04	12.53	3.04	1
4	2.98	12.53	2.98	1
5	2.58	14.19	2.58	1
6	3.67	15.79	3.67	1
7	2.46	12.71	2.46	1
8	3.75	15.53	3.75	1
9	5.63	14.71	5.63	1
10	7.16	10.09	7.16	1
11	11.32	5.30	5.30	2
12	17.55	6.06	6.06	2
13	11.33	5.22	5.22	2
14	16.29	7.57	7.57	2
15	17.78	9.71	9.71	2

Karena hasil dari perhitungan iterasi yang ke-4 pada Tabel 14 nilai centroid tidak berubah atau sama dengan nilai centroid yang sebelumnya maka proses klusterisasi telah selesai. Hasil klusterisasi dari aplikasi weka melakukan iterasi sebanyak 3 kali dengan dengan keseleruhan data sebanyak 366 data dengan fitur *use training set* mendapat hasil visualisasi Gambar berikut.

Tabel 14 Centroid Akhir

Cluster 1	Cluster 2
1	1
2.3	2.4
20.4	24
16.2	19.8
3.7	15.4
2.4	5
1.3	4.6
1.2	1.4

Dapat dilihat pada Gambar 3 hasil dari perluasan data *cluster* berdasarkan anggota yang telah dihitung sesuai jarak data terdekat. Pada Gambar 3 memperlihatkan bahwa *cluster 0* berada di posisi bawah merupakan kelompok data yang memiliki nilai rendah dengan kategori produk x tipe A diperoleh jarak terdekat dengan nilai 4.5, sedangkan *cluster 1* yang berada di posisi atas merupakan kelompok data yang memiliki nilai tinggi dengan kategori produk x tipe B diperoleh jarak terdekat dengan nilai 15.6721 dari atribut merk produk yang sudah diinisialisasikan.



Gambar 3 Grafik Visual Hasil Cluster

Pada Tabel 15 menunjukkan jumlah data responden yang sudah dibagi dalam 2 kelompok yaitu *cluster 0* dan *cluster 1*, kemudian *presentase* jumlah data responden dapat dilihat di setiap *cluster*.

Tabel 15 Hasil Klaster

Cluster	Jumlah	Persen
0	122	33%
1	244	67%

#### 4. Kesimpulan

Berdasarkan hasil dari penelitian, metode *K-Means* dapat digunakan untuk melakukan proses klusterisasi dengan mengubah data menjadi 2 kelompok kemudian kelompok tersebut dapat digunakan untuk mengetahui jenis produk x yang banyak digunakan oleh konsumen. Dari hasil klusterisasi, toko tersebut dapat melakukan pengendalian stok produk x yang banyak dikonsumsi oleh pelanggan agar tidak terjadi *out of stock* berdasarkan kelompok produk x seperti pembahasan dan penjelasan hasil diatas yaitu presentase produk x *cluster 0* di dapat sebesar 33%, sedangkan presentase produk x *cluster 1* di dapat sebesar 67%. Jadi dapat disimpulkan bahwa pengguna produk x tipe B lebih banyak daripada pengguna produk x tipe A. Saran pengembangan untuk penelitian selanjutnya dapat dilakukan dengan melakukan perbandingan menggunakan metode klusterisasi yang lain sebagai pembandingan tingkat akurasi yang lebih akurat dan bisa bermanfaat bagi pihak yang memerlukan.

#### Daftar Pustaka

- [1] R. Andrean, S. Fendy, and A. Nugroho, "Klusterisasi Pengendalian Persediaan Aki Menggunakan Metode K-Means," *JOINTECS (Journal Inf. Technol. Comput. Sci.*, vol. 4, no. 1, pp. 5–12, 2019.
- [2] N. Sapkota, A. Alsadoon, P. W. C. Prasad, A. Elchouemi, and A. K. Singh, "Data Summarization Using Clustering and Classification: Spectral Clustering Combined with k-Means Using NFPH," in *Proceedings of the International Conference on Machine Learning, Big Data, Cloud and Parallel Computing: Trends, Perspectives and Prospects, COMITCon 2019*, 2019, pp. 146–151.
- [3] R. Molaei and S. Alizadeh, "Checking and Analysing Customers Buying Behavior with Clustering Algorithm," *Palma J.*, vol. 16, no. 2, pp. 486–492, 2017.
- [4] R. A. Siregar, "Seleksi Penyerang Utama Menggunakan K-Means Clustering Dan Sistem Pendukung Keputusan Metode Topsis," *Technomedia J.*, vol. 2, no. 1, pp. 37–48, 2017.
- [5] A. Masruro, Kusriani, and E. T. Luthfi, "SISTEM PENUNJANG KEPUTUSAN PENENTUAN LOKASI WISATA MENGGUNAKAN K-MEANS CLUSTERING DAN TOPSIS," *J. Ilm. DASI*, vol. 15, no. 04, pp. 1–5, 2014.
- [6] M. E. Corapcioglu and H. Ogul, "A medical decision support system proposal supported by genomic analysis," in *2017 25th Signal Processing and Communications Applications*

Conference, SIU 2017, 2017.

- [7] R. M. Esteves, T. Hacker, and C. Rong, "Competitive K-means: A new accurate and distributed K-means algorithm for large datasets," in *Proceedings of the International Conference on Cloud Computing Technology and Science, CloudCom*, 2013, vol. 1, pp. 17–24.
- [8] S. Sharma and A. Bhagat, "Data preprocessing algorithm for Web Structure Mining," in *Proceedings on 5th International Conference on Eco-Friendly Computing and Communication Systems, ICECCS 2016*, 2017, pp. 94–98.
- [9] S. Fong, R. Wong, and A. V. Vasilakos, "Accelerated PSO Swarm Search Feature Selection for Data Stream Mining Big Data," *IEEE Trans. Serv. Comput.*, vol. 9, no. 1, pp. 33–45, Jan. 2016.
- [10] S. K. Dwivedi and B. Rawat, "A review paper on data preprocessing: A critical phase in web usage mining process," in *Proceedings of the 2015 International Conference on Green Computing and Internet of Things, ICGCIoT 2015*, 2016, pp. 506–510.
- [11] D. Triyansyah and D. Fitriana, "Analisis Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing," *J. Telekomun. dan Komput.*, vol. 8, no. 3, pp. 163–182, 2018.
- [12] M. Capo, A. Perez, and J. A. A. Lozano, "An efficient Split-Merge re-start for the K-means algorithm," *IEEE Trans. Knowl. Data Eng.*, pp. 1–1, 2020.
- [13] R. Primartha, *BELAJAR MACHINE LEARNING TEORI DAN PRAKTIK*. Informatika Bandung, 2018.
- [14] A. Nugroho and T. M. Fahrudin, "Analisis dan Pemetaan Jumlah Penumpang Kereta Api di Indonesia Menggunakan Metode Statistik Deskriptif dan K-means Clustering," *J. Mantik*, vol. 3, no. 9, pp. 1–9, 2019.
- [15] S. Kapil, M. Chawla, and M. D. Ansari, "On K-means data clustering algorithm with genetic algorithm," in *2016 4th International Conference on Parallel, Distributed and Grid Computing, PDGC 2016*, 2016, pp. 202–206.

**SYSTEMIC: Information System and Informatics Journal**

ISSN: 2460-8092, 2548-6551 (e)

Vol 6 No 1 – Agustus 2020

**Game Promosi Wisata Kota Malang “Kakang Mbakyu” Dengan Menggunakan Decision Tree dan Hierarchy Finite State Machine**Fathurrahman<sup>1</sup>, Yunifa Miftachul Arif<sup>2</sup><sup>1,2</sup> Universitas Islam Negeri Maulana Malik Ibrahim Malang[Fathurrahman637@gmail.com](mailto:Fathurrahman637@gmail.com)<sup>1</sup>, [yunif4@gmail.com](mailto:yunif4@gmail.com)<sup>2</sup>**Kata Kunci***Decision Tree, Finite State Machine, Game, Wisata Kota Malang***Abstrak**

Wisata adalah bepergian secara bersama-sama maupun sendiri dengan tujuan untuk bersenang-senang, menambah pengetahuan dan lain lain. Di zaman sekarang ini wisata sudah mulai menjadi kebutuhan utama bagi setiap manusia, Kota Malang adalah salah satu kota pariwisata dari sekian banyak kota pariwisata di Indonesia yang memiliki banyak destinasi wisata yang beragam, dan juga memiliki daya Tarik tersendiri seperti dari segi iklim, geografis, tradisi dan seni dan budayanya, dengan kelebihan dan keunikan tersebut tidaklah mengherankan apabila Kota Malang terkenal sebagai Kota pariwisata dengan berbagai fasilitas dan pendukung lainnya, Pada penelitian ini Algoritma Decision Tree berhasil terimplementasi pada game dengan menghasilkan gain variable waktu dengan nilai 2,01 dan gain poin 1,86 sehingga variable waktu akan diproses terlebih dahulu sebelum variable poin untuk menghasilkan loncatan perpindahan level sesuai kemampuan player, sedangkan Hierarchy Finite State Machine terbukti berhasil dengan perilaku NPC yang bergerak sudah sesuai dengan rule yang telah di rancang sebelumnya.

**Keywords***Decision Tree, Game, Hierarchy Finite State Machine, Tour, Malang City***Abstract**

Currently, the Tourism Sector in Indonesia is considered the most effective sector contributing to increasing the country's foreign exchange, Foreign exchange earnings were obtained from Indonesian tourism visits which surged and recorded as the highest compared to other countries in Southeast Asia, In particular in the city of Malang there was a significant increase in the number of tourists, based on data listed at [www.malangkota.go.id](http://www.malangkota.go.id) in 2015 the number of tourists entering the city of Malang totaled 3,290,067 people, while 8,265 foreign tourists visited the following year to 3,987,074 for domestic tourists and 9,535 foreign tourists. In this research, Decision Tree algorithm was successfully implemented in the game by producing a time variable gain with a value of 2.01 and a gain point of 1.86 so that the time variable will be processed before the variable points to produce level jumps according to the ability of the player, and for the Hierarchy Finite State Machine, it proved to be successful with the NPC's moving behavior in accordance with the previously designed rules.

**1. Pendahuluan**

Pariwisata (tourism) merupakan salah satu industri besar di dunia, dengan nilai kontribusi terhadap total nilai Produk Domestik Bruto (PDB) dunia yang mencapai 9 persen dan jumlah tenaga kerja yang terserap di industri ini mencapai 200 orang, sektor pariwisata juga mempunyai organisasi dunia yang bernama World Tourism Organization (WTO), Catatan dari World Tourism Organization Pariwisata merupakan sector yang penting WTO mengindikasikan pada tahun 2019

pariwisata Asia Pasifik akan mengalami perkembangan yang menjanjikan terutama dari segi pendapatan (Kurniawan 2013).

Sebagai informasi, jumlah wisatawan mancanegara yang berkunjung ke Indonesia telah menembus 11,5 juta orang, atau naik 10,57% dibanding tahun sebelumnya 10,4 juta. Adapun sepanjang Januari hingga Agustus 2017[1] umlah wisatawan mancanegara yang masuk ke Indonesia tercatat 9,24 juta ([www.cnnindonesia.com](http://www.cnnindonesia.com)) dari jutaan wisatawan

mancanegara pada tahun 2015 sebesar 200.851 kunjungan ke Jawa Timur, namun data terbaru per bulan januari 2017 ke januari 2018 ada perkembangan pertumbuhan wisatawan mancanegara ke Jawa Timur sebesar 54,52%[2].

Tabel 1.1 Perkembangan wisatawan mancanegara ke Jawa Timur

Tahun		Nilai
Januari	2017	17.279
	2018	26.700
Pertumbuhan(%)		54,52

Data tersebut diambil dari Ditjen Imigrasi dan BPS yang sudah diolah kembali oleh asisten deputi industri dan regulasi pariwisata, kementerian pariwisata. Sedangkan berdasarkan data Dinas Budaya dan Pariwisata Kota Malang, selama dua tahun terakhir meningkatkan jumlah wisatawan yang cukup signifikan. Pada tahun 2015 tercatat jumlah wisatawan domestic yang masuk ke kota Malang berjumlah 3.290.067 sedangkan wisman 8.265 pengunjung. Tahun 2016 melonjak menjadi 3.987.074 untuk wisatawan domestik, dan 9.535 orang wisman ([www.Malangkota.go.id](http://www.Malangkota.go.id)).

Melihat peningkatan potensi tersebut, ada beberapa faktor yang harus diperhatikan dan bisa dikerjakan oleh semua kalangan, untuk menjaga konsistensi peningkatan wisatawan. seperti, perbaikan potensi wisata dengan ketersediaan infrastruktur, akses jalan yang memadai dan eksistensi atau kreatifitas promosi harus tetap terus dijalankan maka dibutuhkan strategi untuk memperkenalkan destinasi wisata yang masih banyak belum dikenal baik wisatawan domestik maupun mancanegara, cara yang bisa digunakan yakni seperti dengan promosi media maupun promosi dengan mengadakan event internasional, Menurut kepala Disbudpar Kota Malang, Ida Ayu Made Wahyuni, angka kunjungan wisatawan rata rata meningkat 35 persen setiap tahunnya, untuk itu sejumlah upaya perlu terus dilakukan banyak upaya Pemerintah Kota (Pemkot) Malang mengembangkan aplikasi berbasis Android bernama Malang Menyapa. Aplikasi ini dibuat untuk memperkenalkan pariwisata, yang meliputi tempat wisata, promo paket wisata, dan berita berita terbaru dari kota Malang.

Kasi pemasaran Dinas Kebudayaan dan Pariwisata Kota Malang, Agung H Buana, mengatakan ranah digital harus dimanfaatkan dengan baik sebagai sarana promosi, melalui aplikasi ini, pemkot malang menargetkan ada peningkatan kunjungan wisatawan. terutama kunjungan wisatawan asing. Tahun 2017 target kunjungan wisatawan asing untuk kota malang sebanyak 800.000 hingga 1.5 juta

orang. ([www.beritasatu.com](http://www.beritasatu.com)). Sedangkan promosi dengan event internasional seperti event tahunan yang mampu mendorong jumlah kunjungan wisatawan mancanegara ke kota malang adalah Malang Flower Carnival (MFC). MFC menghadirkan kostum peserta karnaval yang menggunakan bahan baku daur ulang dan didesain menyerupai motif bunga. Menurutnya, penyelenggara MFC 2018 mampu menambah jumlah kunjungan wisatawan mancanegara, termasuk menggerakkan ekonomi Kota Malang, Event terbukti secara efektif dalam meningkatkan kunjungan wisatawan mancanegara kata agung dalam berita online [malang.merdeka.com](http://malang.merdeka.com) ([www.malang.merdeka.com](http://www.malang.merdeka.com)).

Pemerintah Kota Malang telah menyiapkan kalender kegiatan wisata atau Calendar of event 2019, yang diharapkan mampu menjadi acuan bagi wisatawan dalam negeri maupun mancanegara untuk berkunjung ke Kota Malang. Dari sekian langkah yang di lakukan oleh pemerintah penulis belum menemukan adanya upaya promosi dengan menggunakan Game edukasi Game edukasi adalah salah satu jenis media yang digunakan untuk memberikan pengajaran, menambah pengetahuan penggunanya melalui suatu media unik dan menarik, Game edukatif adalah permainan yang dibuat dan dirancang untuk merancang daya pikir termasuk meningkatkan konsentrasi dan memecahkan masalah (Handriyantini,2009) Dari uraian diatas dapat disimpulkan Game edukasi adalah salah satu bentuk Game yang dapat berguna untuk menunjang proses belajar mengajar secara lebih menyenangkan dan lebih kreatif.

Pada khususnya Game umumnya memiliki beberapa komponen yang penting yaitu skenario (alur cerita), Level (tingkatan), skor (nilai) karakter, dan *obstacle*/ (rintangan), dari sekian banyak komponen ini, pada penelitian ini peneliti akan fokus meneliti tentang level (tingkatan) dan NPC atau musuh terdapat sebuah tingkatan kesulitan, dimulai dari tingkat kesulitan yang berawal dari tingkat kesulitan mudah (*easy*), sedang (*normal*) dan sulit (*hard*) [3] perubahan tingkat kesulitan ditandai dengan adanya penambahan level di dalam Game. Penambahan level di dalam Game digunakan agar Game lebih menarik dikarenakan mempunyai tingkat kesulitan yang mengasah kemampuan dan terpacu untuk berpikir dan menyelesaikan level satu ke level yang lain dengan intensitas yang bertahap dan teratur sesuai kemampuan player, pada penelitian kali ini penulis akan meneliti metode dalam mengatur perpindahan level dengan perolehan poin dan waktu yang menjadi indikator perpindahan level dalam Game dan juga metode untuk mendesain perilaku NPC. Sama halnya dengan level, NPC juga digunakan agar Game lebih menarik pada penelitian kali ini, peneliti akan membuat Game edukasi dengan tujuan

untuk ikut mendukung pertumbuhan yang signifikan terhadap pariwisata yang ada di kota Malang yang berjudul “Kakang dan Mbakyu Kota Malang” Kakang dan Mbakyu adalah julukan untuk duta di kota Malang, dalam hal ini duta pariwisata,

Game ini membantu memperkenalkan destinasi apa saja yang ada di Malang dan berisi konten-konten edukatif yang ditujukan sesuai fungsi yang sebelumnya dijelaskan, namun agar Game lebih menarik maka diperlukan kecerdasan buatan yang membuat Game menjadi lebih berkualitas, dalam Game ini penulis menggunakan kecerdasan buatan dalam hal ini Hierarchy Finite State Machine (HFMSM), untuk mendesain perilaku NPC, karena menurut peneliti HFMSM ini bisa mengatur lebih detail apa yang akan dilakukan NPC sesuai keadaan yang sedang berjalan, sedangkan Decision Tree digunakan untuk mengatur perpindahan level dalam Game dimana menggunakan parameter kemampuan player dalam hal ini seberapa cepat dan seberapa banyak player bisa mengumpulkan poin dan menghabiskan waktu dalam menyelesaikan permainan.

## 2. Metode Penelitian

### 2.1 Landasan Teori

#### 2.1.1 Game

Pengertian game jika diartikan secara bahasa (berasal dari bahasa Inggris) adalah permainan. Dengan bantuan teknologi, kini game dapat diartikan secara lebih luas lagi. Jadi, secara istilah pengertian game adalah permainan yang diprogram pada suatu perangkat yang dapat dijalankan secara offline maupun online.[4].

#### 2.1.2 Unity 3D

Unity 3D merupakan sebuah Game engine, yaitu software pengolah gambar, grafik, suara, input, dan lain-lain yang ditujukan untuk membuat Game. Unity3D merupakan Game engine multiplatform, yang mampu di-publish secara standalone(.exe) berbasis web, android, ios, XBOX, maupun PS3, dengan catatan mendapatkan lisensi. Versi gratis Cuma bisa di-publish ke bentuk standalone dan Web [5].

#### 2.1.3 Decision Tree

Decision Tree adalah struktur flowchart yang menyerupai Tree (pohon), dimana setiap simpul internal menandakan suatu tes pada atribut, setiap cabang merepresentasikan hasil tes, dan simpul daun merepresentasikan kelas atau distribusi kelas. Alur pada Decision Tree ditelusuri dari simpul akar ke simpul daun yang memegang prediksi.[10].

#### 2.1.4 Entropy dan information gain

Sebuah obyek yang diklasifikasikan dalam

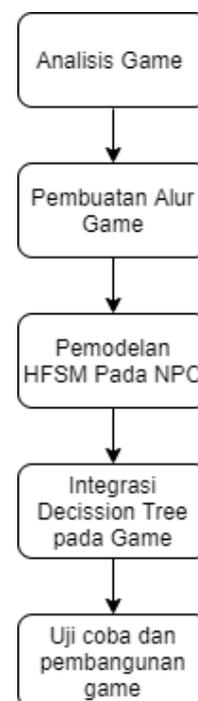
pohon harus di tes nilai entropy nya, entropy adalah ukuran dari teori informasi yang dapat mengetahui karakteristik dari impurity dan homogeneity dari kumpulan data, dari nilai entropy tersebut kemudian dihitung nilai information gain (IG) masing-masing atribut. Entropy (S) merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ dan -) dari sejumlah data acak pada ruang sampel S, Entropy dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas [7].

### 2.1.5 Hierarchy Finite State Machine

Hierarchy Finite State machine (HFMSM) adalah sebuah algoritma yang dikembangkan dari finite state machine(FSM) sama seperti FSM, HFMSM memiliki beberapa factor untuk menjalankan kerja system, yaitu state, event, transition dan action yang membedakan adalah pada HFMSM terdapat super state atau clustering dan juga sub state, clustering menggabungkan beberapa state terpisah menjadi satu state [9].

## 2.2 Analisis Game

Pada tahap ini peneliti menganalisis game yang akan dibangun guna merancang apa saja yang dibutuhkan dalam membangun sebuah game yang sesuai dengan tujuan awal yakni bisa memberikan edukasi sekaligus mempromosikan destinasi wisata yang ada. Dalam penelitian ini penulis melakukan beberapa metodologi dalam proses membangun game memperoleh data atau informasi, yang bertujuan untuk menyelesaikan permasalahan, dengan digambarkan dalam bentuk kerangka pikir, sebagaimana Gambar 1.



Gambar 1. Kerangka pikir

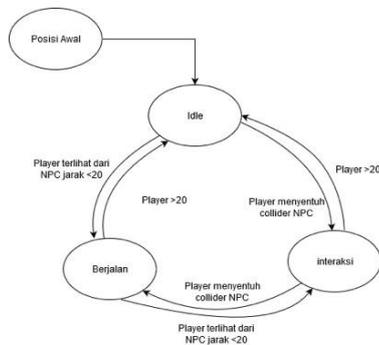
### 2.2.1 Pembuatan alur game

Dalam proses ini akan dilakukan perancangan baik dari storyboard, bahasa pemrograman, pengumpulan dan pembuatan kebutuhan audio visual, perancangan kebutuhan Game (status, barang dalam Game, dsb), dan segala sesuatu yang akan dibutuhkan, dalam game ini bahasa pemrograman yang digunakan yaitu C#, untuk kebutuhan game yang lain seperti asset, banyak mengambil dari asset store dan beberapa juga membuat asset sendiri.

### 2.2.2 Pemodelan HFSM pada NPC

FSM adalah sebuah metodologi perancangan system control yang menggambarkan tingkah laku atau prinsip kerja system dengan menggunakan tiga hal berikut : state, event, dan action, [8] pada kali ini Setelah proses sebelumnya, masuk ke pemodelan HFSM NPC pada game, Pada tahap ini, sebelum mendesain HFSM, terlebih dahulu mendesain FSM dari masing masing NPC, karena HFSM adalah representasi dari fsm fsm yang ada di dalam fsm, untuk game ini masing masing NPC mempunyai FSM yang berbeda, berikut adalah FSM tiap NPC

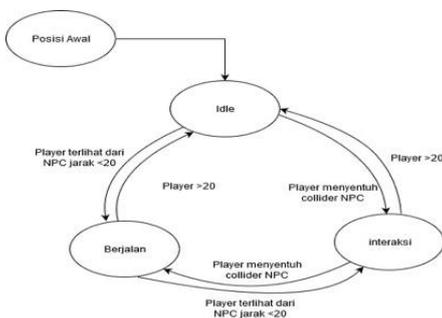
#### ● FSM NPC Tourleader



Gambar 3. FSM NPC Tourleader

FSM NPC *Tourleader* merupakan gambar NPC *Tourleader* akan bekerja, pada NPC *Tourleader* ini akan digambarkan dengan dua NPC sesuai dengan *Hierarchy* yang sudah digambarkan sebelumnya yakni ketika NPC bertemu *player* dan ketika NPC akan memberikan quiz terkait.

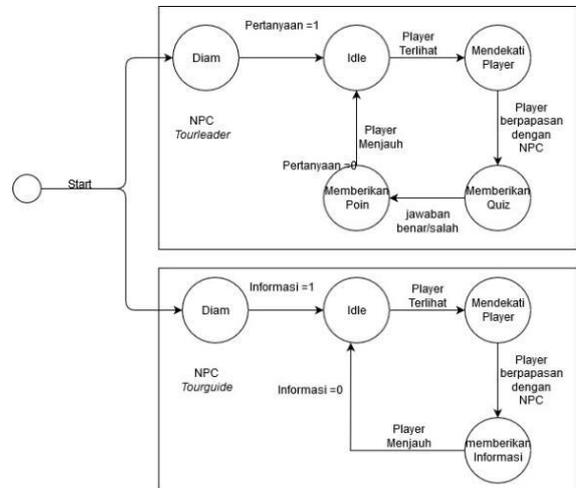
#### ● FSM NPC Tourguide



Gambar 4. FSM NPC Tourguide

FSM NPC *Tourguide* merupakan desain NPC *Tourguide* akan bekerja, pada NPC *Tourguide* ini akan digambarkan gerakan NPC sesuai dengan *Hierarchy* yang sudah digambarkan sebelumnya yakni ketika NPC bertemu *player* dan ketika NPC akan memberikan pertanyaan terkait. NPC *Tourguide* bertemu *player*

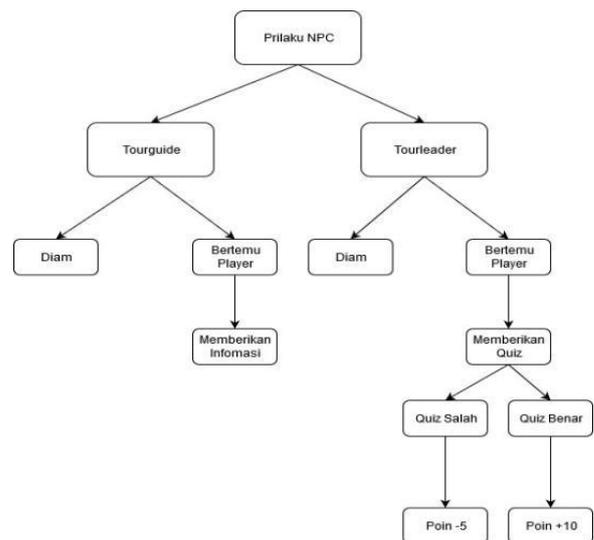
#### ● Top Level FSM NPC



Gambar 5. Top level FSM NPC

#### ● Hierarchy Finite State Machine

Setelah FSM dan Top level FSM didesain, dan dijadikan hirarki, seperti inilah bentuk representasi dari hierarchy finite state machine NPC pada game ini.

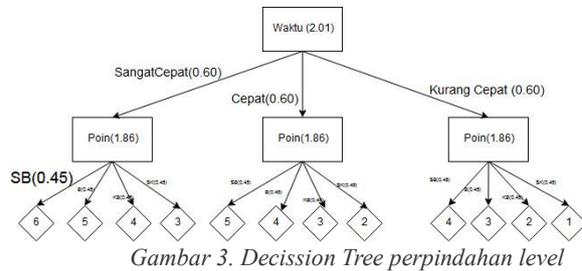


Gambar 3. FSM NPC Tourleader

### 2.4 Integrasi Decision tree

Pada tahap ini dilakukan pengintegrasian antara Game kakang mbakyu pariwisata kota Malang dengan algoritma decision tree yang digunakan untuk perpindahan level atau dari scene satu ke scene yang lainnya, integrasi metode dilakukan dengan prosedur decision tree yaitu perhitungan

entropy dan information gain, setelah perhitungan berulang ulang sampai semua atribut pohon memiliki kelas dan tidak bisa lagi dilakukan proses perhitungan[6] metode decision tree diletakkan di kotak rahasa yang menjadi finish di setiap levelnya, dalam perpindahan level tersebut metode ini digunakan untuk mengklarifikasi data, dan data yang diklarifikasi dijadikan inputan untuk menjadi paramter pengguna dalam perpindahan dari level yang sedang dimainkan ke level yang lainnya.



Gambar 3. Decission Tree perpindahan level

Gambar di atas adalah pohon keputusan yang dihasilkan dari penelitian ini, gambar ini menjelaskan bahwa berdasarkan nilai gain yang dihasilkan dari masing masing kategori, kategori waktu memiliki nilai gain yang lebih besar dari pada kategori point sehingga dalam proses perpindahan level, kategori waktu lah yang di proses terlebih dahulu,

**2.5 Uji coba dan pembangunan game**

Dalam proses ini dilakukan uji coba untuk mengetahui implementasi metode apakah sudah berhasil atau belum, uji coba kali ini terbagi atas dua, yaitu uji coba HFSM dan Uji coba perpindahan level menggunakan decision tree. Setelah dilakukan pengujian HFSM, pergerakan NPC yang didapatkan sudah sesuai dengan rancangan HFSM yang sudah di rancang.

Tabel 1 Pengujian HFSM

No.	User	NPC	Keterangan
1	Jarak player NPC <i>Tourguide</i> <20 & >10	Berbalik kearah player dan berjalan mendekati player	Sesuai
2	Player tepat didepan NPC <i>Tourguide</i> dengan jarak <10	Memberikan informasi seputar destinasi	Sesuai
3	Jarak player NPC <i>Tourleader</i> <20 & >10	Berbalik kearah player dan berjalan mendekati player	Sesuai
4	Player tepat didepan NPC <i>Tourleader</i> dengan jarak <10	Memberikan quiz seputar informasi yang didapatkan	Sesuai
5	Player menjawab benar quiz <i>Tourleader</i>	Memberikan poin ke player 10 Poin	Sesuai
6	Player menjawab salah quiz <i>Tourleader</i>	Mengurangi poin player sebesar 5 poin	Sesuai
7	Player berada jauh dari NPC	NPC bergerak <i>default</i> seperti awal	Sesuai

Dari hasil pengujian, perpindahan scene yang dihasilkan dari rule *Decission Tree* bahwa waktu yang dihabiskan player lah terlebih dahulu yang dihitug dari pada point dikarenakan nilai gain waktu lebih besar dari pada nilai gain point, sehingga semakin cepat player menyelesaikan permainan maka semakin tinggi loncatan perpindahan scene yang diperoleh player, dan untuk perpindahan scene yang di uji sudah sesuai dengan rule *Decission tree* yang dihasilkan.

Tabel 2 Pengujian Decission Tree

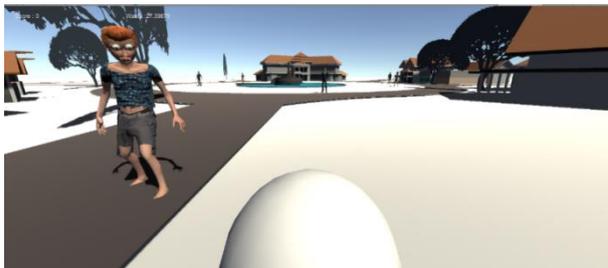
No.	Input		Output	Keterangan
	Waktu	Poin	Akses Scene	
1	Sangat cepat	Sangat Banyak	6	Sesuai
2	Sangat cepat	Banyak	5	Sesuai
3	Sangat cepat	Kurang banyak	4	Sesuai
4	Sangat cepat	Sangat kurang	3	Sesuai
5	Cepat	Sangat banyak	5	Sesuai
6	Cepat	banyak	4	Sesuai
7	Cepat	Kurang banyak	3	Sesuai
8	Cepat	Sangat kurang	2	Sesuai
9	Kurang cepat	Sangat banyak	4	Sesuai
10	Kurang cepat	Banyak	3	Sesuai
11	Kurang cepat	Kurang banyak	2	Sesuai
12	Kurang cepat	Sangat kurang	1	Sesuai

Berdasarkan Tabel 1 dan 2 di atas terlihat bahwa algoritma *Decission Tree* berhasil mengatur perpindahan level dalam Game ini, dengan menggunakan variable score (sangat banyak, banyak, kurang banyak, sangat kurang) dan variable waktu (sangat cepat, cepat dan kurang cepat). *Hierarchy Finite State Machine* berhasil diimplementasikan dalam mendesain perilaku pada NPC, metode ini cocok diaplikasikan pada game ini, sesuai dengan rancangan yang sudah dibuat, bentuk dari representasi metode ini berupa diagram yang didalamnya terdapat FSM yang menggambarkan tingkah laku NPC oleh system control ke dalam bentuk hirarki yang sederhana dan mudah untuk diimplementasikan.

### 3. Hasil Dan Pembahasan

#### 3.1 Tampilan Game

Berikut adalah tampilan tampilan pada game yang berhasil di bangun



Gambar 4 Tampilan awal game

Dalam game ini pemain langsung di letakkan di scene atau level pertama yaitu destinasi balai kota atau tugu kota malang, yang masuk ke kategori mudah, player bertugas untuk berkelan atau jalan jalan mengitari destinasi dengan menemui beberapa NPC dalam hal ini ada NPC tourleader dan NPC Tourguide.



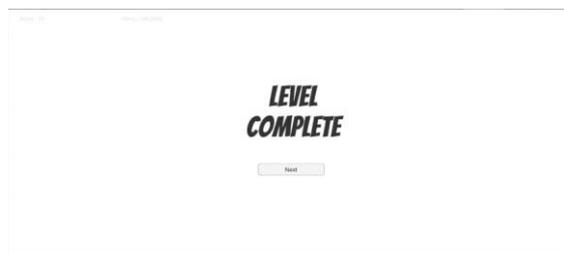
Gambar 5 tampilan NPC Tourguide bertemu player

Gambar diatas adalah ketika player bertemu dengan NPC Tourguide ketika TOurguide memberikan informasi terkait destinasi wisata yang sedang dikunjungi.



Gambar 6 tampilan NPC Tourleader bertemu player

Gambar diatas adalah ketika player bertemu dengan NPC Tourleader ketika Tourleader memberikanquiz terkait destinasi wisata yang sedang dikunjungi.



Gambar 7 Tampilan game level berakhir

Gambar diatas adalah tampilan game ketika player berhasil menyelesaikan satu scene, dan pada gambar ini algoritma decision tree di proses, ketika player menekan tombol next, maka algoritma decision tree mulai berjalan memproses nilai inputan yang telah diperoleh player, dan player akan diarahkan ke level sesuai dengan hasil kemampuan player.

#### 3.2 Kesimpulan

Berdasarkan hasil dari implementasi dan pengujian yang dilakukan peneliti maka dapat ditarik kesimpulan sebagai berikut:

- 1) Algoritma Decision Tree berhasil mengatur perpindahan level dalam Game ini, dengan menggunakan variable score (sangat banyak, banyak, kurang banyak, sangat kurang) dan variable waktu (sangat cepat, cepat dan kurang cepat). Pada penelitian ini algoritma decision tree menghasilkan gain variable waktu dengan nilai 2,01 dan dan gain poin 1,86 sehingga variable waktu akan diproses terlebih dahulu sebelum variable poin.
- 2) Hierarchy Finite State Machine berhasil diimplementasikan dalam mendesain perilaku pada NPC, metode ini cocok diaplikasikan pada game ini, sesuai dengan rancangan yang sudah dibuat, bentuk dari representasi metode ini berupa diagram yang didalamnya terdapat FSM yang menggambarkan tingkah laku NPC oleh system control ke dalam bentuk hirarki yang sederhana dan mudah dipahami, sehingga memudahkan untuk diimplementasikan ke dalam Bahasa pemrograman, secara garis besar dalam hirarki yang dimaksud terbagi menjadi dua bagian yaitu FSM Tourguide dan FSM Tourleader, dalam percobaan ini NPC berperilaku sudah sesuai dengan state yang ada pada diagram.

#### Daftar Pustaka

[1] Utama Devisa. CNN Indoneia Diambil dari <http://www.cnnindonesia.com>

[2] Badan Pusat Statistik Jawa Timur , 2016. “Statistik Pariwisata Jawa Timur

[3] Dewi Intan. 2016. Penggunaan Decision Tree untuk Game pemilihan jalur sepeda. Skripsi UIN MALANG

- [4] Santoso, Erick , pembuatan game dengan menerapkan Metode Decission Tree : UCB 1 untuk menentukan pemilihan strategi dalam AI
- [5] Andi, “Mudah membuat game 3 Dimensi,” 2014.
- [6] Eka Pandu chyntia, Edi Ismanto, “Metode Decission tree algoritma C.45 dalam mengklasifikasi data penjualan bisnis gerai makanan cepat saji”, Jurnal riset system informasi dan Teknik informatika Vol.3, 2018
- [7] Rani.L.N”Klasifikasi nasabah menggunakan algoritma C.45 sebagai dasar pemberian kredit” Jurnal inovtek polbeng – seri informatika Vol.1 No.2, 2016
- [8] Setiawan Iwan, “Perancangan Software embedded system berbasis FSM”, UNDIP, 2006.
- [9] Mayovio Ahmad, “Penerapan Decission Making NPC dengan metode Hierarchical Finite State Machine pada 2D endless runner game”, Jurnal pengembangan teknologi informasi dan ilmu computer, 2019.
- [10] Zalilia, L. “Penerapan data mining untuk IDS”. Tugas akhir mata kuliah EC7010 program studi Teknik elektro Institut Teknologi Bandung. 2007.



**SYSTEMIC**  
INFORMATION SYSTEM AND INFORMATICS JOURNAL



9 772460 809004