

**Protected by PDF Anti-Copy Free**

**(Upgrade to Pro Version to Remove the Watermark)**  
**METODE *HYBIRD* DALAM PENGELOMPOKKAN**  
**KEMAMPUAN *CRISTUNG* SISWA BERBASIS**  
***MAC* *LEARNING***



**SKRIPSI**

**Diajukan Sebagai Syarat Untuk Menyelesaikan Pendidikan**  
**Program Sarjana ( S-1)**  
**Pada Program Studi Informatika**

**Oleh :**  
**AMANDA SALSABILA**  
**NIM : 2102020001**

**PROGRAM STUDI INFORMATIKA**  
**FAKULTAS ILMU TEKNIK**  
**UNIVERSITAS BINA INSAN LUBUKLINGGAU**  
**2025**

**Protected by PDF Anti-Copy Free**

**(Upgrade to Pro Version to Remove the Watermark)**

**HALAMAN PENGESAHAN SKRIPSI**



**TEKNIK INFORMATIKA UNIVERSITAS BINA INSAN  
METODE *HYBIRD* DALAM PENGELOMPOKKAN  
KEMAMPUAN CALISTUNG SISWA BERBASIS  
*MACHINE LEARNING***

**Oleh:**

**AMANDA SALSABILA**

**NIM : 2102020001**

**Pembimbing I**

**(Andri Anto Tri Susilo, M.Kom)**

**Lubuklinggau, 23 Januari 2025  
Pembimbing II**

**(Nelly Khairani Daulay, M.Kom)**

**Mengesahkan  
Dekan Fakultas Ilmu Teknik  
Universitas Bina Insan,**

**(Dr. Rudi Kurniawan, S.T, M.Kom)**

# Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

## HALAMAN PERSETUJUAN TIM PENGUJI SKRIPSI



Pada hari Kamis tanggal 23 bulan Januari tahun 2025 telah dilaksanakan Sidang Skripsi oleh Program Studi Informatika Fakultas Ilmu Teknik Universitas Bina Insan.

Nama : Amanda Salsabila  
NIM : 2102020001  
Judul Skripsi : Metode *Hybird* Dalam Pengelompokkan Kemampuan Calistung Siswa Berbasis *Machine Learning*.

### Komisi Penguji

1. Ketua : **Andri Anto Tri Susilo, M.Kom** ( ..... )
2. Sekretaris : **Nelly Khairani Daulay, M.Kom** ( ..... )
3. Anggota : **Armanto, M.Kom** ( ..... )

**Mengetahui,  
Kepala Program Studi Informatika  
Fakultas Ilmu Teknik  
Universitas Bina Insan**

**(Budi Santoso, M.Kom)**

## HALAMAN MOTTO DAN PERSEMBAHAN



### *MOTTO*

- ♥ “ Hujan itu tidak pernah menatangimu, dia hanya melewatimu, karena hidup adalah silih bergantinya musim, maka terimalah dan tetaplah bertahan, dia pasti akan berlalu “. (Ustadz Hanan Attaki)
- ♥ “ Sedari kecil dibentuk untuk mejadi mesin penghancur badai, maka tak pantas diri ini tumbang hanya karna perkataan seseorang “.

### *Skripsi ini dipersembahkan untuk :*

- ♥ Cinta pertama dan pintu surgaku, ayahanda Sugiarno dan ibunda Suhartini yang sangat amanda cintai. Terima kaih atas cinta, kasih sayang, segala do'a dan dukungan kalian. Amanda dapat bertahan dalam setiap proses dan bisa menyelesaikan skripsi ini, semoga sehat selalu aaminn.
- ♥ Untuk kakak-kakak yang amanda banggakan. terimakasih atas segala motivasi, dan dukungan, serta mengajarkan untuk terus berusaha tiada henti untuk menjadi orang yag sukses, selalu mendo'akan dan memberi semangat yang tiada henti.
- ♥ Untuk diri sendiri Amanda Salsabila. Terimakasih sudah bekerja keras, bertahan dan berjuang sejauh ini, menghadapi berbagai hal sulit dan Bahagia dalam hidup. Tak pernah meyerah sesuliat apapun skripsi ini dengan melewati dan menyelesaikan sebaik dan semaksimal mungkin, Aku bangga kamu hebat, kamu keren amanda.

**Protected by PDF Anti-Copy Free**  
**(Upgrade to Pro Version to Remove the Watermark)**  
**HALAMAN PERNYATAAN**



Saya yang bertanda tangan di bawah ini .

Nama Mahasiswa/i : Amanda Salsabila

NIM : 2102020001

Program Studi : Informatika

Fakultas : Ilmu Teknik

Menyatakan dengan sesungguhnya bahwa penelitian dan penulisan Skripsi yang saya susun sebagai persyaratan untuk memperoleh gelar Sarjana (S-1) Universitas Bina Insan, merupakan hasil kerja saya sendiri dan tidak menyuruh orang lain yang mengerjakannya. Ada pun bagian tertentu dalam penulisan skripsi ini yang saya kutip dari hasil karya orang lain dan telah saya tuliskan sumbernya secara jelas sesuai dengan norma, kaidah dan etika penulisan ilmiah.

Jika dikemudian hari ternyata terbukti bahwa penelitian dan tugas akhir ini bukan hasil kerja saya sendiri atau plagiat dalam bagian-bagian tertentu, maka saya bersedia dikenakan sanksi sesuai dengan peraturan perundangan yang berlaku.

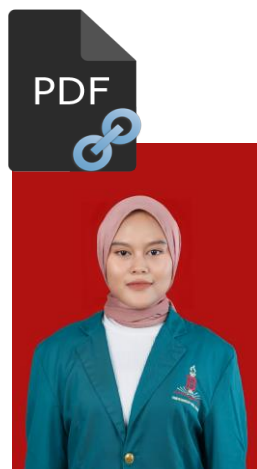
**Lubuklinggau, 23 Januari 2025**  
**Penulis,**

**Amanda Salsabila**  
**NIM 2102020001**

**Protected by PDF Anti-Copy Free**

**(Upgrade to Pro Version to Remove the Watermark)**

## **DAFTAR RIWAYAT HIDUP**



### **Biodata**

Nama : Amanda Salsabila  
Tempat / Tanggal Lahir : Lubuklinggau, 06 Februari 2003  
Jenis Kelamin : Perempuan  
Agama : Islam  
Alamat : Jl. Nangka Kacung No.77 Rt.03  
Kel.Ponorogo kec. Lubuklinggau Utara II


### **Pendidikan**

- SD : SD Negeri 48 Kota Lubuklinggau
- SMP/MTS Sederajat : SMP Negeri 8 Kota Lubuklinggau
- SMA/MAN/SMK Sederajat : SMA Negeri 3 Kota Lubuklinggau

# Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

## ABSTRACT

 Calistung skills (reading, writing, and arithmetic) are essential foundations in the academic development of elementary school students. This study aims to cluster students' calistung abilities using a hybrid method based on machine learning. The data used consists of the reading, writing, and calculating scores of students from two elementary schools in Lubuklinggau City. The hybrid method applied in this research combines the K-Means Clustering algorithm for initial grouping with the K-Nearest Neighbors (KNN) algorithm for the classification process. The analysis process includes data preprocessing, implementation of the K-Means algorithm, validation of clusters, and classification using K-Nearest Neighbors (KNN) to ensure the accuracy of the clustering results. This study shows that clustering using K-Means successfully groups students based on their scores, which can be used for further analysis in education, resulting in 3 score clusters: Medium (0), Low (1), and High (2). Subsequently, the K-Nearest Neighbors (KNN) classification model was tested with  $k=1$ , which achieved the highest accuracy of 95%, providing excellent accuracy in the K-Nearest Neighbors (KNN) classification model with an accuracy of 97%. The precision for cluster 0 is 97%, for cluster 1 is 100%, and for cluster 2 is 96%. The recall for cluster 0 is 97%, for cluster 1 is 93%, and for cluster 2 is 99%. The F1-Score for cluster 0 is 97%, for cluster 1 is 96%, and for cluster 2 is 98% in classifying students' data categorized as "very good."

**Keywords:** *Hybrid Method, K-Means Clustering, K-Nearest Neighbors (KNN), Calistung Skills*

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### ABSTRAK

Kemampuan calistung (membaca, menulis, dan berhitung) merupakan dasar penting dalam perkembangan akademik siswa sekolah dasar. Penelitian ini bertujuan untuk mengelompokkan kemampuan calistung siswa dengan menggunakan metode *hybrid* berbasis *machine learning*. Data yang digunakan berupa nilai kemampuan membaca, menulis, dan berhitung siswa dari dua Sekolah Dasar di Kota Lubuklinggau. Metode hybrid yang diterapkan dalam penelitian ini menggabungkan algoritma *K-Means Clustering* untuk pengelompokan awal dengan algoritma *K-Nearest Neighbors (KNN)* untuk proses klasifikasi. Proses analisis meliputi tahap preprocessing data, implementasi algoritma *K-Means*, validasi kluster, serta klasifikasi menggunakan *K-Nearest Neighbors (KNN)* untuk memastikan akurasi hasil pengelompokan. Penelitian ini menunjukkan bahwa *clustering* menggunakan *K-Means* berhasil mengelompokkan siswa berdasarkan nilai mereka, yang dapat digunakan untuk analisis lebih lanjut dalam pendidikan, menghasilkan 3 kluster nilai yaitu Menengah(0), Rendah (1), dan Tinggi (2). Setelah itu, pada pengujian model klasifikasi *K-Nearest Neighbors (KNN)* dengan  $k=3$  yang memiliki akurasi tertinggi yaitu 95% memberikan akurasi pada pengujian model klasifikasi *K-Nearest Neighbors (KNN)* yang sangat baik dengan akurasi 97%, *Precision cluster 0* sebesar 97%, *cluster 1* sebesar 100%, *cluster 2* sebesar 96%. *Recall cluster 0* sebesar 97%, *cluster 1* sebesar 93%, *cluster 2* sebesar 99%. *F1-Score cluster 0* sebesar 97%, *cluster 1* sebesar 96%, *cluster 2* sebesar 98% dalam klasifikasi data siswa yang dikategorikan “sangat baik” .

**Kata kunci:** Metode Hybrid, K-Means Clustering, K-Nearest Neighbors (KNN), Kemampuan Calistung Siswa

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### KATA PENGANTAR



Alhamdulillah puji dan penulis ucapkan kepada Allah SWT atas segala rahmat dan karunia-Nya yang telah memberikan kekuatan dan kesempatan, sehingga penulis dapat menyelesaikan Skripsi ini dengan maksimal, Untuk diajukan sebagai syarat menyelesaikan pendidikan program Sarjana (S-1) Pada Program Studi Informatika Fakultas Ilmu Teknik Universitas Bina Insan. Sholawat beserta salam semoga tetap tercurahkan kepada bagi Nabi Muhammad SAW, keluarga, sahabat, serta umatnya hingga akhir zaman.

Selama proses penulisan dan penyusunan Skripsi ini, penulis telah berusaha sebaik-baiknya untuk dapat menyelesaikan Skripsi ini baik tepat pada waktunya. Penulis menyadari bahwa Skripsi ini tentunya masih jauh dari sempurna dan mungkin terdapat kesalahan baik sengaja maupun tidak sengaja. Oleh karena itu, kritik dan saran yang membangun tentunya sangat diharapkan dari berbagai pihak.

Penulis mengucapkan banyak terima kasih kepada pihak-pihak yang telah membantu selama proses penyelesaian skripsi ini diantaranya yaitu:

1. Orang Tua dan Keluargaku Tercinta yang telah banyak memberikan do'a, dukungan, dan bantuannya dalam penulisan Skripsi ini.
2. Bapak Dr. H. Sardiyono, M.M. selaku Rektor Universitas Bina Insan.
3. Bapak Dr. Muhamad Akbar, S.T., M.IT selaku Wakil Rektor I Universitas Bina Insan.
4. Bapak Wakhid Nur Mukhlis, M.Pd., M.M selaku Wakil Rektor II Universitas Bina Insan.
5. Bapak Dr. Rudi Kurniawan, S.T., M.Kom. selaku Dekan Fakultas Ilmu Teknik Universitas Bina Insan yang telah banyak memberikan bimbingan dan arah dalam penulisan skripsi ini.
6. Bapak Budi Santoso, S.Kom., M.kom. selaku Kepala Program Studi Informatika Fakultas Ilmu Teknik Universitas Bina Insan yang telah banyak memberikan bimbingan dan arah dalam penulisan Skripsi ini.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

7. Bapak Andri Anto Tri Susilo, M.Kom. selaku Pembimbing I yang telah banyak memberikan bimbingan dan arahan dalam penulisan Skripsi ini.
8. Ibu Nelly Khairani Daulay, M.Pd. selaku Pembimbing II yang telah banyak memberikan bimbingan dan arahan dalam penulisan Skripsi ini.
9. Bapak/Ibu selaku Penguji yang telah banyak memberikan bimbingan dan arah dalam penulisan Skripsi ini.
10. Bapak Kepala Sekolah dan Staf Dewan Guru SD Negeri 48 Kota Lubuklinggau
11. Ibu Kepala Sekolah dan Staf Dewan Guru SD Negeri 23 Kota Lubuklinggau.
12. Seluruh Staf Dosen dan Karyawan Universitas Bina Insan Lubuklinggau yang telah banyak memberikan ilmu pengetahuan dan bimbingan kepada penulis.

Akhir kata semoga penelitian ini dapat bermanfaat bagi untuk penelitian selanjutnya.

Lubuklinggau, 23 Januari 2025  
**Penulis**

**Amanda Salsabila**  
**(NIM 2102020001)**

# Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

## DAFTAR ISI



HALAMAN PENGESAHAN	ii
HALAMAN PERSETUJUAN TIM PENGUJI SKRIPSI	iii
HALAMAN MOTO DAN PERSEMBAHAN	iv
HALAMAN PERNYATAAN	v
DAFTAR RIWAYAT HIDUP	vi
ABSTRACT	vii
ABSTRAK	viii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
<b>BAB I PENDAHULUAN</b>	
1.1. Latar Belakang Penelitian	1
1.2. Identifikasi Masalah	3
1.3. Rumusan Masalah	3
1.4. Batasan Masalah	3
1.5. Tujuan dan Manfaat Penelitian	4
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1. Literatur	7
2.2. Penelitian Terdahulu yang Relevan	12
2.3. Kerangka Berfikir	16
<b>BAB III METODELOGI PENELITIAN</b>	
3.1. Analisa Sistem	18
3.1.1. Analisa Sistem Yang Berjalan	18
3.1.2. Alternatif Pemecahan Masalah	18
3.1.3. Metode Analisa	19

## Protected by PDF Anti-Copy Free

<b>(Upgrade to Pro Version to Remove the Watermark)</b>	
3.2. Teknik Pemilihan Informan (Populasi dan Sampel) .....	21
3.2.3. Teknik Pengumpulan Data .....	21
3.2.3. Perancangan Instrumen .....	33
3.2.3. Teknik Analisa .....	34
3.2.3.1. Pengolahan Data .....	34
3.2.3.2. Pengujian Model .....	35
3.3. Tempat dan Waktu Penelitian .....	37
3.3.3. Alat dan Bahan .....	37

### **BAB IV HASIL PENELITIAN DAN PEMBAHASAN**

4.1 Gambaran Umum .....	39
4.1.1. Gambaran Umum (Tempat Penelitian) .....	39
4.1.2. Struktur Organisasi .....	40
4.2 Hasil Penelitian .....	41
4.3 Pembahasan .....	51
4.1.3. Penerapan Metode Analisa dan Validasi data .....	51
4.1.4. Pengujian Hasil Analisa .....	54

### **BAB V KESIMPULAN DAN SARAN**

5.1 Kesimpulan .....	71
5.1 Saran .....	71

### **DAFTAR PUSTAKA**

### **LAMPIRAN**

**Protected by PDF Anti-Copy Free**

**(Upgrade to Pro Version to Remove the Watermark)**

**DAFTAR TABEL**



	<b>Halaman</b>
Tabel 2. 1 Perhitungan Algoritma K-means Clustering.....	11
Tabel 2. 2 Penelitian Terdahulu.....	12
Tabel 3. 1 Sampel Data Penelitian .....	21
Tabel 3. 2 Waktu Penelitian .....	37
Tabel 4. 2 Confusion Matrix .....	46

# Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

## DAFTAR GAMBAR



	<b>Halaman</b>
Gambar 2. 1 Kerangka Berfikir .....	17
Gambar 3. 1 Analisis Konsep Penelitian.....	19
Gambar 4.1 Struktur Organisasi Sekolah Dasar Negeri 48 Kota Lubuklinggau.....	40
Gambar 4. 2 Struktur Organisasi Sekolah Dasar Negeri 23 Kota Lubuklinggau.....	41
Gambar 4. 3 Hasil Data Cluster Dengan K-Means .....	43
Gambar 4. 4 Grafik Metode Elbow Menentukan Nilai K Optimal .....	43
Gambar 4. 5 Hasil Pengujian Model Prediksi Cluster Pada Data Baru.....	44
Gambar 4. 6 Grafik Akurasi K-Nearest Neighbors (KNN) Nilai K Terbaik .....	45
Gambar 4. 7 Grafik Confusion Matrix .....	45
Gambar 4. 8 Hasil Akurasi, Precision, Recall dan F1-Score.....	49
Gambar 4. 9 Hasil Inputan Akhir .....	50
Gambar 4. 10 Import Library .....	51
Gambar 4. 11 Read_CSV .....	51
Gambar 4. 12 Data Describe .....	52
Gambar 4. 13 Informasi Dataset .....	52
Gambar 4. 14 Membagi Data Fitur dan Target .....	53
Gambar 4. 15 Pembagian Data Training dan Testing.....	53
Gambar 4. 16 Hasil Clustering.....	54
Gambar 4. 17 Clustering dengan K Optimal.....	55
Gambar 4. 18 Rata Rata dan Menentukan Centeroid.....	56
Gambar 4. 19 Centeroid Akhir .....	57
Gambar 4. 20 Jumlah Optimal Cluster Menggunakan Metode Elbow .....	58
Gambar 4. 21 Visualisasi 3D Hasil Clustering.....	59
Gambar 4. 22 Mengakses data terbaru csv dan Penggabungan file .....	60
Gambar 4. 23 Label Baru dari Cluster .....	60
Gambar 4. 24 Penghapusan Kolom yang Bernilai String .....	61
Gambar 4. 25 Pebagian Data dan inisialisasi Encode .....	62
Gambar 4. 26 Model K-Nearest Neighbors (KNN) .....	62
Gambar 4. 27 Model K-Nearest Neighbors (KNN) Prediksi Cluster Untuk Data Baru.....	63
Gambar 4. 28 Data Sampel Untuk Pengujian Prediksi Cluster Pada Data Baru .....	63
Gambar 4. 29 Hasil Prediksi Cluster Dengan Data Baru .....	64
Gambar 4. 30 Pengecekan Nilai Nan Data Training dan Mengisi Nilai Hilang.....	65
Gambar 4. 31 Nilai Akurasi K1-10 .....	65
Gambar 4. 32 Grafik Accurasi K-Nearest Neighbors (KNN) Pada Nilai K.....	66
Gambar 4. 33 Menampilkan Nilai K Terbaik.....	66
Gambar 4. 34 Model Confusion Matrix dan Accuracy Total .....	66
Gambar 4. 35 Visualisasi Grafik Confusion Matrix.....	67
Gambar 4. 36 Akurasi, Precision, Recal Sensitivity, Specificity, FPR, FNR, dan F1-score. ....	68
Gambar 4. 37 Menampilkan Accuracy, Precision, Recall dan F1-score .....	68
Gambar 4. 38 Data Terbaru Hasil Klasifikasi .....	69
Gambar 4. 39 Hasil Input Akhir.....	70

# Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

## DAFTAR LAMPIRAN



Lampiran 1. Judul Penelitian ...	77
Lampiran 2. Surat Permohonan Izin Penelitian .....	78
Lampiran 3. Surat Permohonan Izin Penelitian .....	79
Lampiran 4. Surat Balasan Permohonan Izin Pelitian .....	80
Lampiran 5. Surat Balasan Permohonan Izin Pelitian .....	81
Lampiran 6. Lembar Bimbingan Proposal Dosen Pembimbing I.....	82
Lampiran 7. Lembar Bimbingan Proposal Dosem Pembimbing II .....	83
Lampiran 8. Lembar Bimbingan Proposal Dosem Pembimbing II .....	84
Lampiran 9. Lembar Perbaikan Seminar Proposal Skripsi.....	85
Lampiran 10. Lembar Bimbingan Skripsi Pembimbing I.....	86
Lampiran 11. Lembar Bimbingan Skripsi Pembimbing II .....	87
Lampiran 12. Dokumentasi Bersama Kepala Sekolah SD .....	88
Lampiran 13. Dokumetasi Pengambilan Data Nilai Siswa.....	88

**Protected by PDF Anti-Copy Free**  
**(Upgrade to Pro Version to Remove the Watermark)**  
**BAB I PENDAHULUAN**

### 1.1. Latar Belakang Penelitian



Di era digital abad ke-21, pendidikan mengalami perubahan signifikan dengan penerapan teknologi pintar, analisis data, dan kecerdasan buatan, yang menciptakan pengalaman belajar lebih personal dan kreatif serta memperluas akses informasi[1]. Pemerintah dan masyarakat perlu mempertimbangkan pendidikan dalam meningkatkan kualitas sumber daya manusia di tengah persaingan globalisasi yang semakin ketat[2]. Selain itu, globalisasi mendorong perkembangan literasi dan numerasi, termasuk kemampuan berpikir dan menalar yang mendasari pendidikan dasar seperti membaca, menulis, dan berhitung (calistung) menjadi fondasi pendidikan di sekolah dasar[3]. Membaca melibatkan pengenalan huruf, kata, serta pemahaman makna; menulis merupakan komunikasi tidak langsung yang produktif dan ekspresif; sementara berhitung berkaitan dengan penggunaan pola dan aturan untuk menyelesaikan masalah[4]. Keterampilan membaca, menulis, dan berhitung (calistung) adalah dasar penting bagi setiap individu, membantu mengenali huruf dan angka serta memfasilitasi komunikasi melalui bahasa dan simbol[5]. Kemampuan ini berpengaruh positif pada perkembangan bahasa dan logika seseorang. Bagi anak-anak usia sekolah dasar, calistung mendukung perkembangan bahasa dan logika, dengan membaca dan menulis meningkatkan pemahaman informasi dan berhitung memperkuat pemikiran logis[6].

Penelitian ini berfokus pada dua Sekolah Dasar Negeri di Kota Lubuklinggau, yaitu SD Negeri 48 dan SD Negeri 23. Kedua sekolah ini terletak di area perkotaan Lubuklinggau. Dari pengamatan awal, ditemukan bahwa sekolah-sekolah ini menghadapi tantangan dalam mengatasi perbedaan kemampuan calistung siswa di dalam kelas. Para guru sering kali melihat adanya perbedaan signifikan dalam kemampuan membaca, menulis, dan berhitung di antara siswa yang berada dalam satu kelas. Hal ini menyebabkan proses pengajaran menjadi kurang efektif, karena guru kesulitan memberikan perhatian yang tepat kepada masing-masing siswa.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Mengelompokkan siswa berdasarkan kemampuan calistung secara manual memerlukan banyak waktu dan tenaga, terutama di kelas besar, dan sering dilakukan secara manual tanpa data akurat[7]. Dalam konteks Education 4.0, yang menekankan integrasi teknologi cerdas dan data besar, penerapan *machine learning* menjadi semakin penting[8]. Dengan kemajuan teknologi dan *machine learning*, metode seperti K-Means Clustering dapat mempermudah pengelompokan siswa dengan akurasi lebih tinggi[9]. Metode *K-Means Clustering* ini digunakan karena dapat membagi siswa ke dalam beberapa kelompok berdasarkan kemiripan dalam kemampuan mereka[10]. Metode hybrid yang menggabungkan algoritma *machine learning*, yaitu clustering dan classification, memungkinkan pengelompokan kemampuan siswa dilakukan dengan lebih efektif, cepat, dan akurat. Metode ini semakin populer karena meningkatkan performa dalam prediksi dan pengelompokan, serta dapat mengidentifikasi pola kemampuan siswa secara mendalam dengan memanfaatkan data nilai baca, tulis, dan hitung [11]. Penerapan metode ini juga memudahkan guru dalam memberikan intervensi pembelajaran yang lebih sesuai dengan tingkat kemampuan siswa berdasarkan kelompok[12].

Untuk mengatasi masalah tersebut, dibutuhkan suatu sistem yang mampu mengklasifikasi kemampuan membaca, menulis, dan berhitung (calistung) siswa sesuai dengan karakteristik yang dimilikinya. Dengan adanya penelitian ini, diharapkan dapat bermanfaat dalam membantu pihak sekolah, khususnya di SD Negeri 48 dan SD Negeri 23 Kota Lubuklinggau, dalam menentukan kemampuan siswa dan memberikan intervensi yang tepat. Sistem klasifikasi ini dapat membantu guru untuk lebih mudah mengelompokkan siswa berdasarkan kemampuan mereka dalam calistung, sehingga proses pembelajaran dapat dioptimalkan sesuai dengan kebutuhan setiap siswa. Penelitian ini menggunakan metode *hybrid* dengan memanfaatkan algoritma *K-Means Clustering* untuk mengelompokkan kemampuan siswa ke dalam beberapa cluster berdasarkan nilai membaca, menulis, dan berhitung. Setelah proses pengelompokan, metode klasifikasi berbasis *machine learning* akan digunakan untuk memperkirakan kemampuan siswa baru atau siswa yang

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

belum diklasifikasi sebelumnya. Dengan demikian, penelitian ini dapat memberikan solusi yang lebih baik dalam proses evaluasi akademik siswa serta membantu pihak sekolah dalam menyusun kebijakan pendidikan yang lebih efisien dan tepat sasaran. Penelitian ini difokuskan pada pengelompokan dan klasifikasi kemampuan membaca, menulis, dan berhitung siswa di Sekolah Dasar Kota Lubuklinggau dengan menggunakan metode *K-Means Clustering* dan klasifikasi *K-Nearest Neighbors (KNN)*. Yang berjudul “**Metode Hybrid dalam Pengelompokan Kemampuan Calistung Siswa Berbasis Machine Learning**”. Data yang digunakan adalah nilai Membaca, Menulis, dan Menghitung siswa. Dari hasil evaluasi yang diperoleh sekolah. Diharapkan penelitian ini dapat memberikan manfaat praktis dalam upaya meningkatkan kualitas pembelajaran di tingkat sekolah dasar.

### 1.2. Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan, penulis mengidentifikasi masalah utama dalam pengelompokan kemampuan calistung siswa di Sekolah Dasar 48 dan 23 kota Lubuklinggau yaitu, dalam proses pengelompokan kemampuan calistung siswa masih dilakukan secara manual melalui observasi dan penilaian sederhana. Metode ini bersifat subjektif dan tidak konsisten, yang sering kali mengakibatkan pengelompokan yang tidak akurat dan tidak efektif.

### 1.3. Rumusan Masalah

Berdasarkan identifikasi masalah maka rumusan masalah yang dibahas dalam penelitian ini yaitu, bagaimana agar proses pengelompokan dan klasifikasi kemampuan calistung siswa tidak lagi dilakukan secara manual dengan metode yang dilakukan dapat memengaruhi akurasi yang akurat dan efektif?

### 1.4. Batasan Masalah

Untuk memberikan arah yang lebih terfokus pada permasalahan penelitian ini, pembatasan dilakukan sebagai berikut:

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

- a. Data penelitian yang digunakan ini adalah data nilai calistung siswa dari dua Sekolah Dasar yaitu, SD Negeri 48 Lubuklinggau dan SD Negeri 23 Lubuklinggau.
- b. Jumlah sampel data yang digunakan sebanyak 508 data siswa kelas 1 sampai kelas 6.
- c. Data yang terkumpul adalah data dari tahun (2023-2024).
- d. Penelitian ini hanya fokus pada pengelompokan kemampuan calistung, yaitu pada nilai membaca, menulis, dan berhitung.
- e. Metode yang akan digunakan dalam penelitian ini terbatas pada *K-Means* untuk pengelompokan dan *K-Nearest Neighbors (KNN)* untuk klasifikasi. Metode lain dalam machine learning tidak akan dipertimbangkan dalam penelitian ini.
- f. Data yang digunakan dalam penelitian ini diambil dari hasil tes calistung yang telah dilakukan sebelumnya oleh siswa, tanpa melakukan pengujian tambahan atau pengumpulan data baru selama penelitian berlangsung.

### 1.5. Tujuan dan Manfaat Penelitian

Tujuan dan manfaat pada penelitian ini akan dibahas pada penjelasan dibawah ini :

#### 1.5.1. Tujuan Penelitian

Berdasarkan rumusan masalah sebelumnya maka tujuan dari penelitian yang dilakukan penulis adalah untuk:

- a. Penelitian ini bertujuan untuk mengidentifikasi variasi kemampuan calistung siswa di Sekolah Dasar dan mengelompokkan mereka berdasarkan hasil tes menggunakan metode *K-Means*. Mengevaluasi keefektifan model *hybrid* yang dikembangkan dalam memetakan kemampuan siswa.
- b. Menerapkan metode *K-Nearest Neighbors (KNN)* untuk mengklasifikasikan siswa baru berdasarkan pengelompokan yang telah dilakukan, sehingga memudahkan dalam penentuan strategi pembelajaran yang sesuai.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

- c. Penelitian ini bertujuan untuk membangun model pengelompokan dan klasifikasi kemampuan calistung siswa yang dapat digunakan oleh guru untuk merancang strategi pembelajaran yang lebih efektif.
- d. Penelitian ini juga bertujuan untuk memberikan rekomendasi bagi guru dan pihak sekolah dalam meningkatkan proses pembelajaran berdasarkan hasil pengelompokan dan klasifikasi yang diperoleh.

### 1.5.2. Manfaat Penelitian

Manfaat dari penelitian ini mencakup pada “Metode *Hybrid* dalam Pengelompokan Kemampuan Calistung Siswa Berbasis *Machine Learning*”. Bagi penulis, penelitian ini memiliki sejumlah manfaat antar lain:

1. Bagi Pendidikan penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan metode evaluasi dan pengelompokan siswa yang lebih efektif, sehingga dapat meningkatkan kualitas pembelajaran di Sekolah Dasar.
2. Bagi Guru hasil penelitian ini akan memberikan informasi yang berguna bagi guru dalam memahami karakteristik kemampuan calistung siswa, sehingga mereka dapat merancang dan menerapkan strategi pembelajaran yang lebih sesuai dengan kebutuhan masing-masing kelompok siswa.
3. Bagi Siswa dengan penerapan pengelompokan yang tepat, siswa diharapkan dapat menerima pembelajaran yang lebih terfokus dan sesuai dengan tingkat kemampuan mereka, sehingga dapat meningkatkan hasil belajar dan motivasi belajar.
4. Bagi Peneliti Selanjutnya penelitian ini dapat menjadi dasar bagi penelitian lebih lanjut dalam bidang pendidikan dan teknologi, khususnya dalam penerapan machine learning untuk analisis data pendidikan.

### 1.6. Sistematika Penulisan

#### **BAB I : PENDAHULUAN**

Dalam bab ini berisikan mengenai Latar Belakang, Identifikasi Masalah, Rumusan Masalah, Batasan

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Masalah, Tujuan dan Manfaat Penelitian, serta Sistem dalam Penulisan.

**BAB II : KAJI PUSTAKA**

Dalam bab ini berisikan teori yang mendasari masalah yang diteliti yang berhubungan dengan tema penelitian.

**BAB III : METODOLOGI PENELITIAN**

Dalam bab ini berisikan metode penelitian, metode pengumpulan data, metode analisis, metode pengujian dan pengolahan data.

**BAB IV : HASIL DAN PEMBAHASAN**

Dalam bab ini berisikan tentang implementasi dan simulasi pengolahan data ke dalam bahasa pemrograman.

**BAB V : KESIMPULAN DAN SARAN**

Dalam bab ini berisikan tentang kesimpulan dari seluruh penelitian skripsi yang telah dilakukan dan saran-saran / masukan-masukan yang berguna dimasa yang akan datang.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### BAB II TINJAUAN PUSTAKA



#### 2.1. Literatur

##### 2.1.1. Pengertian Membaca, Menulis, dan Berhitung (Calistung)

Menurut Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional, pasal 4 ayat 5, menyatakan bahwa pendidikan bertujuan membudayakan keterampilan membaca, menulis, dan berhitung (calistung) di masyarakat[13]. Kemampuan calistung ini penting sebagai dasar pembelajaran dan keterampilan hidup, sehingga perlu dikuasai sejak dini. Kompetensi ini diperkenalkan sejak jenjang TK/PAUD untuk mendukung program literasi dan numerasi, terutama bagi siswa sekolah dasar kelas rendah[14]. Kemampuan membaca, menulis, dan berhitung merupakan tiga keterampilan dasar yang harus dikuasai siswa pada fase pendidikan dasar. **Membaca** adalah kemampuan memahami teks tertulis, termasuk bunyi, makna, dan struktur kalimat, untuk mendapatkan informasi secara tepat[15]. **Menulis** adalah proses menuangkan ide atau informasi secara jelas dan terstruktur dalam bentuk tulisan, meliputi kemampuan menyusun kalimat secara logis. [16]. **Berhitung** mencakup literasi numerasi dan kemampuan melakukan operasi dasar seperti penjumlahan, pengurangan, perkalian, dan pembagian. [17].

##### 2.1.2. Machine Learning

*Machine learning* adalah cabang dari kecerdasan buatan (*Artificial Intelligence*) yang memungkinkan sistem komputer untuk belajar dari data tanpa perlu diprogram secara eksplisit[18]. Metode Machine Learning (ML) dalam pendidikan membantu tenaga pengajar menganalisis pola data kompleks, memberikan wawasan tentang efektivitas pembelajaran, peningkatan keterampilan siswa, dan evaluasi kurikulum[19]. Dalam *machine learning*, algoritma dilatih dengan sejumlah data dan kemudian digunakan untuk membuat prediksi atau keputusan berdasarkan pola yang ditemukan dari data tersebut. Berdasarkan jenisnya, machine learning dapat dibagi menjadi empat kategori utama: *supervised learning*, *unsupervised*

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

*learning, Semi-Supervised Learning dan Reinforcement Learning* [20].

### 1) *Supervised Learning*

*Supervised learning* melibatkan data yang telah dilabeli, di mana model dilatih untuk menghasilkan output berdasarkan input yang sudah memiliki label[21].

### 2) *Unsupervised Learning*

*Unsupervised Learning* digunakan untuk menemukan pola tersembunyi dalam data yang belum dilabeli. Algoritma *K-Means Clustering* adalah salah satu metode *unsupervised learning* yang paling populer[22].

### 3) *Semi-supervised Learning*

*Semi-supervised Learning* menggabungkan unsur-unsur dari *supervised* dan *unsupervised learning*, di mana sebagian data memiliki label, sementara sebagian lainnya tidak dilabeli[22].

### 4) *Reinforcement Learning*

*Reinforcement Learning* model pembelajaran ini menggunakan sistem reward dan penalti, dimana model belajar mendapatkan reward dan berusaha menghindari penalti.[23].

### 2.1.3. *Data Mining*

*Data Mining* adalah proses untuk mengidentifikasi pola, tren, dan informasi tersembunyi dari data dalam jumlah besar menggunakan teknik analisis yang kompleks[24]. *Data mining* adalah proses semi-otomatis yang memanfaatkan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstrak serta mengidentifikasi pengetahuan potensial yang bermanfaat dari database besar. *Data mining* merupakan bagian dari proses *KDD (Knowledge Discovery in Databases)*, yang melibatkan tahapan-tahapan seperti pemilihan data, prapengolahan, transformasi, data mining, dan evaluasi hasil. Secara umum, *KDD* juga dikenal sebagai pangkalan data[25]. *Data mining* mencakup berbagai metode yang bertujuan untuk menemukan pola yang sebelumnya tidak diketahui dalam data yang telah terkumpul. Proses ini bertujuan untuk menghasilkan kesimpulan atau

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

fakta baru dengan memfilter data guna menemukan pola atau anomali. Data mining memiliki lima jenis utama, yaitu *classification*, *clustering*, *association*, dan *sequence mining*[25].

### 2.1.4. Clustering/Klasterisasi

*Clustering* merupakan proses mengelompokkan sejumlah objek fisik atau abstrak ke dalam kategori yang serupa[26]. Clustering adalah proses pengelompokan objek berdasarkan hubungan antar objek dengan tujuan memaksimalkan kesamaan dalam satu kelompok dan meminimalkan kesamaan antar kelompok. Karakteristik setiap cluster muncul dari kesamaan atribut data yang tergabung di dalamnya, tanpa ditentukan sebelumnya[27]. Tujuannya adalah menemukan klaster yang berkualitas tinggi dalam waktu yang efisien. Teknik ini bisa digunakan, misalnya, untuk menentukan lokasi optimal titik akses berdasarkan posisi pengguna[28].

Penerapan teknik clustering dapat dimulai dengan menentukan topik yang akan dibahas, lalu menggunakan kotak, garis, atau lingkaran untuk menandai ide-ide yang muncul. Ide-ide tersebut dapat dituliskan dalam bentuk kata atau frasa[29].

### 2.1.5. Algoritma $K$ – Means Clustering

*K-means* merupakan salah satu algoritma dari *clustering*, *K-Means Clustering* adalah metode *unsupervised learning* yang mengelompokkan data ke dalam beberapa klaster berdasarkan kesamaan karakteristik. Algoritma ini dimulai dengan memilih sejumlah titik pusat (*centroid*) secara acak, kemudian data dikelompokkan ke dalam klaster yang memiliki jarak paling dekat dengan *centroid*[30].

*K-Means* adalah algoritma yang digunakan untuk mempartisi guna memisahkan data ke dalam kelompok yang berbeda. Algoritma ini sangat sederhana dalam hal clustering. Proses *clustering* dari algoritma ini adalah dengan membagi atau mengelompokkan kumpulan data menjadi *cluster-cluster*. Pada metode *K-means*, data dibagi menjadi beberapa kelompok, dimana setiap kelompok mempunyai ciri-ciri yang mirip atau identik dengan kelompok lainnya, namun kelompok yang lain mempunyai ciri-ciri yang

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

berbeda[31].

Dalam melakukan inisialisasi, metode *K-means* memiliki beberapa langkah, yaitu sebagai berikut :

- 1) Pilih  $K$  sebagai jumlah
- 2) Inisialisasi  $K$  *centroid* secara acak, Hitung jarak setiap titik data ke *centroid* menggunakan rumus *Euclidean Distance*. Penginisialisasian ini dapat dilakukan dengan beberapa cara. Cara yang paling sering digunakan dalam melakukan penginisialisasian ini dengan cara acak (random) sehingga pusat-pusat *cluster* diberi nilai awal dengan angka-angka acak. Untuk setiap titik data  $x$ , kita perlu menghitung jarak antara titik data dengan setiap *centroid*. Rumus yang paling umum digunakan untuk menghitung jarak adalah *Euclidean Distance*:

$$d(x, c) = \sqrt{\sum_{i=1}^n (x_i - c_i)^2} \quad (1)$$

Keterangan :

$d(x, c)$  = jarak antara titik data  $x$  dan centroid  $c$ .

$x$  = titik data (koordinat atau nilai fitur dari data).

$c$  = centroid (koordinat pusat cluster).

$x_i$  = nilai fitur ke- $i$  dari titik data  $x$ .

$c_i$  = nilai fitur ke- $i$  dari centroid  $c$ .

$n$  = jumlah fitur atau dimensi data.

Fungsi utama dari rumus ini adalah menghitung jarak sehingga data dapat dikelompokkan ke centroid yang memiliki jarak terdekat.

- 3) Kelompokkan titik data ke centroid terdekat.
- 4) Perbarui posisi centroid menggunakan rumus Mean Update. Setelah titik-titik data dikelompokkan berdasarkan kedekatannya dengan centroid, posisi centroid akan diperbarui menjadi rata-rata dari semua titik data di dalam cluster. Rumus pembaruan centroid terlihat dibawah ini :

$$\mu_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i \quad (2)$$

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Keterangan :

$\mu_j$  = posisi centroid berdasarkan cluster  $j$ .

$c_j$  = sekumpulan titik yang bergabung dalam cluster ke- $j$ .

$x_i$  = titik data ke- $i$  di cluster tersebut.

$|c_j|$  = jumlah total titik data dalam cluster ke- $j$ .

Fungsi utama dari rumus ini adalah memperbarui posisi centroid berdasarkan rata-rata data yang sudah bergabung dalam cluster, sehingga posisi centroid lebih mencerminkan data di cluster tersebut.

- 5) Ulangi langkah 2 hingga 4 sampai centroid tidak berubah atau perubahan kecil (konvergensi tercapai).

Perhitungan *K-Means Clustering* dapat dilihat pada *Table 2.1*

*Perhitungan algoritma K-Means Clustering.*

Tabel 2. 1 Perhitungan Algoritma K-Means Clustering

No	Nama	Membaca	Menulis	Menghitung	C1	C2	C3	JARAK TERDEKAT	CLUSTER	JARAK TERDEKAT <sup>2</sup>
1	Alifa syakia	75	80	78	4.718613	26.17463	5.830952	4.718612733	C1	22.26530612
2	Atika putri	80	80	75	8.601139	26.97736	0	0	C3	0
3	Kayra nasya	75	80	79	5.409742	26.87213	6.403124	5.40974178	C1	29.26530612
4	Boy candra	60	65	60	23.65785	3.333333	29.15476	3.333333333	C2	11.11111111
5	Cinta larasati	65	75	70	9.202928	14.24001	16.58312	9.202927662	C1	84.69387755
6	Dafa ananda	70	80	79	4.30709	25.10865	10.77033	4.307089552	C1	18.55102041
7	Dedek noelo	75	80	79	5.409742	26.87213	6.403124	5.40974178	C1	29.26530612
8	Delisa sabila	70	80	70	5.248907	19.22094	11.18034	5.248906592	C1	27.55102041
9	Denjaka halim	60	65	60	23.65785	3.333333	29.15476	3.333333333	C2	11.11111111
10	Diki saputra	80	80	70	9.948849	24.55153	5	5	C3	25
11	Dimas nabil	70	65	60	20.76349	6.666667	23.45208	6.666666667	C2	44.44444444
12	Dio deandra	70	80	70	5.248907	19.22094	11.18034	5.248906592	C1	27.55102041
13	Diva salsabila	80	80	75	8.601139	26.97736	0	0	C3	0
14	Dorra innasya	80	80	76	8.659076	27.54592	1	1	C3	1
15	Elvira syakira	80	80	79	9.485757	29.3901	4	4	C3	16
										347.8095238

### 2.1.6. Metode Hybrid

*Hybrid* adalah proses yang menggabungkan berbagai pendekatan. Misalnya penggabungan metode machine learning seperti cluster dan klasifikasi[33].

Metode *hybrid* dalam konteks *Machine Learning* adalah kombinasi yang sering kali dimaksud untuk meningkatkan performa model dengan menggabungkan kekuatan dari dua atau lebih algoritma yang berbeda.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Dalam penelitian ini, *Metode Hybrid* yang digunakan adalah perpaduan antara *K-Means Clustering* dan metode klasifikasi menggunakan algoritma *KNN* [34]. Dengan menggabungkan kedua metode ini, tidak hanya mampu mengelompokkan data berdasarkan kesamaan kemampuan calistung mereka, tetapi juga memprediksi atau mengklasifikasikan data baru yang belum terkelompokkan berdasarkan pola yang dihasilkan dari pengelompokan sebelumnya. *Metode Hybrid* yang menggabungkan clustering dan klasifikasi dapat memberikan hasil yang lebih akurat dalam pengelompokan dan prediksi. Algoritma *K-Means* digunakan untuk memisahkan siswa ke dalam beberapa cluster, kemudian algoritma klasifikasi *K-Nearest Neighbors (KNN)* digunakan untuk mengevaluasi model tersebut dan membuat prediksi tentang perkembangan siswa[35].

### 2.2. Penelitian Terdahulu yang Relevan

Beberapa penelitian yang relevan atau terdahulu dalam penelitian ini menjadi salah satu acuan penulis dalam melakukan penelitian sehingga penulis dapat memperkaya teori yang digunakan dalam mengkaji penelitian yang dilakukan. Berikut adalah Penelitian Terdahulu yang Relevan yang dilakukan oleh beberapa peneliti sebelumnya terlihat pada *Table 2.2 Penelitian Terdahulu*.

Tabel 2. 2 Penelitian Terdahulu

No.	Nama Peneliti	Judul	Tahun	Metode	Hasil
1.	Fadil Danu Rahman, Mulki Indana Zulfa, Acep Taryana	Custering dan klasifikasi data cuaca kota cilacap menggunakan k-means dan random forest	2024	K-Means clustering dan algoritma Random Forest	Melalui penggunaan algoritma <i>Random Forest</i> , model klasifikasi berhasil memprediksi kondisi cuaca dengan tingkat

## Protected by PDF Anti-Copy Free

[\(Upgrade to Pro Version to Remove the Watermark\)](#)



akurasi yang memuaskan. Meskipun demikian, penurunan performa pada rentang tahun 2018-2023 menunjukkan adanya tantangan dalam memodelkan pola cuaca yang kompleks. Analisis

- |    |  |  |      |                                    |   |
|----|--|--|------|------------------------------------|---|
| 2. | Adella Putri Riani, Apriade Voutama, Taufik Ridwan | Penerapan K-Means Clustering Dalam Pengelompokan Hasil Belajar Peserta Didik Dengan Metode Elbow | 2023 | algoritma K-Means dan metode Elbow | Hasil perbandingan nilai Davies Bouldin Index (DBI) juga membuktikan bahwa $c=3$ merupakan jumlah cluster yang paling ideal karena nilai DBI yang paling mendekati 0 dibandingkan dengan jumlah |
|----|--|--|------|------------------------------------|---|

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



cluster lain. Hasil akurasi menggunakan DBI pada penelitian ini adalah 88,4%.

- |    |   |   |      |                                 |   |
|----|---|---|------|---------------------------------|---|
| 3. | Trovina Tige Kati, Reynaldi Thimotius Abineno, Arini Aha Pekuwali | Penerapan K-Means Clustering Untuk Mengelompokkan Performa Siswa Dalam Pelajaran Bahasa Indonesia | 2024 | algortma k-means                | Hasil penelitian ini menunjukkan bahwa penggunaan algortma k-means dapat membantu mengelompokkan siswa dengan nilai yang tinggi dan yang rendah ke dalam kelompok yang lebih kecil. |
| 4. | Mey Susi Munthea, Sri Lestari Rahayu                              | Klasifikasi Data Siswa Yang Layak Dapat Bantuan Menggunakan Metode K-Nearest Neighbors            | 2024 | Metode K-Nearest Neighbor (KNN) | <i>K-Nearest Neighbor (KNN)</i> merupakan proses normalisasi karena metode ini mengalikan hasil penilaian setiap atribut. Hasil perkalian tersebut                                  |

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



belum bermakna jika belum dibandingkan (dibagi) dengan nilai standar. Bobot untuk atribut manfaat berfungsi sebagai pangkat positif dalam proses perkalian, sementara bobot biaya berfungsi sebagai pangkat negative.

- |    |  |  |      |   |  |
|----|--|--|------|---|--|
| 4. | Dewi Eka Putri,<br>Eka Praja Wiyata<br>Mandala | Hybrid Data Mining berdasarkan Klasterisasi Produk untuk Klasifikasi Penjualan | 2024 | algoritma K-Means dan K-Nearest Neighbors | Hasil pengujian diperoleh klasifikasi produk baru adalah penjualan tinggi dengan ketepatan dalam klasifikasi mencapai 85.7143 %. Penggunaan metode hybrid antara algoritma K-Means dan K-Nearest Neighbors |
|----|--|--|------|---|--|
-

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



memberikan pengaruh yang signifikan dalam menentukan hasil klasifikasi

---

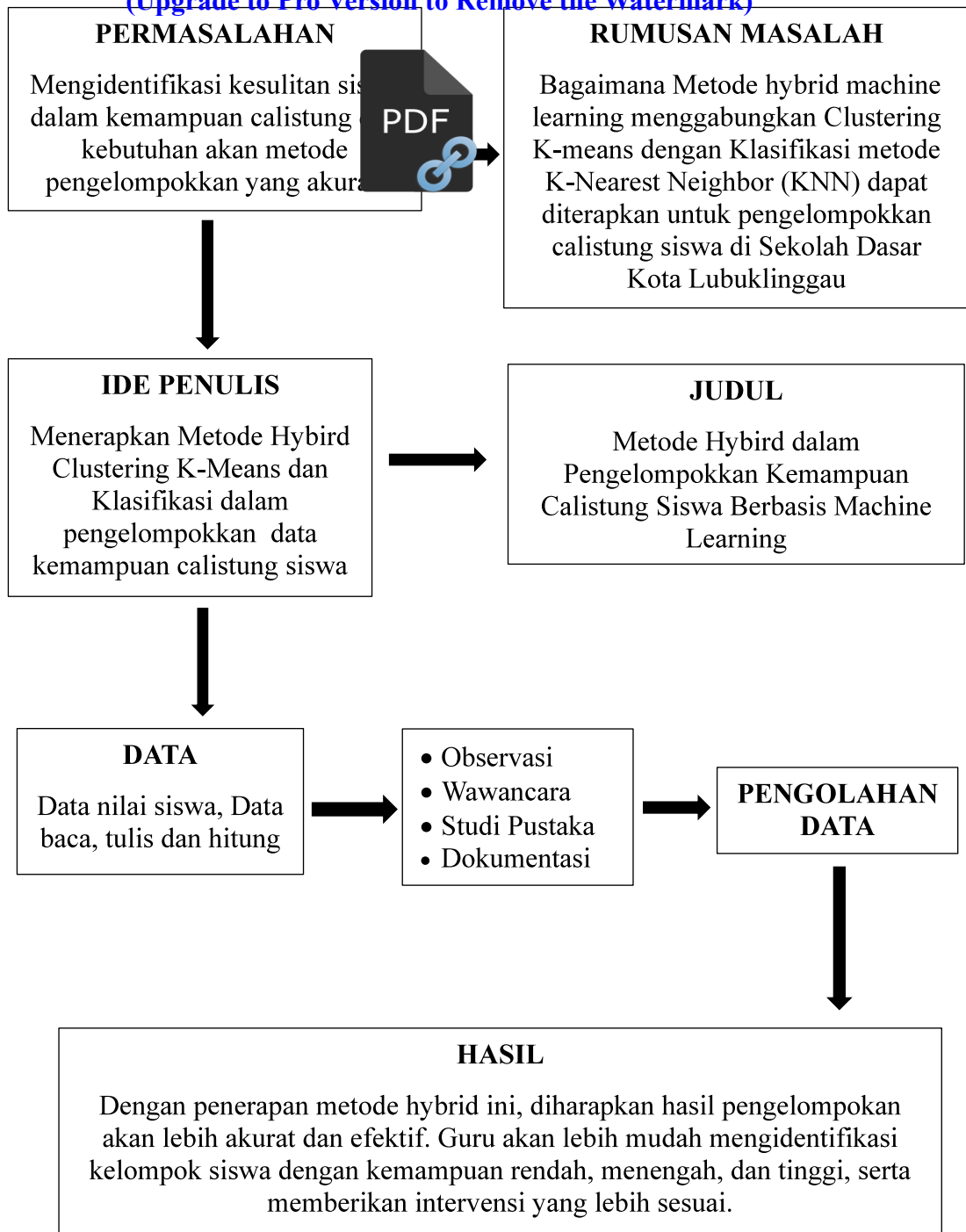
Perbedaan dengan penelitian terkait sebelumnya adalah penelitian ini akan dilakukan dalam 2 (dua) tahap yaitu metode klustering dan klasifikasi. Perbedaan selanjutnya adalah dari sumber data yang akan digunakan dalam penelitian yaitu penelitian ini menggunakan data nilai calistung siswa yang diperoleh dari dua Sekolah Dasar, yaitu SD Negeri 48 kota lubuklinggau dan SD Negeri 23 kota lubuklinggau.

### 2.3. Kerangka Berfikir

Kerangka berpikir dalam penelitian ini disusun untuk menggambarkan alur logis dari proses penelitian yang akan dilakukan. Diawali dengan identifikasi permasalahan terkait kemampuan dasar membaca, menulis, dan berhitung (calistung) siswa yang sering kali sulit dinilai secara akurat, penelitian ini menawarkan solusi berupa penggunaan metode hybrid berbasis *machine learning*. Pendekatan ini menggabungkan teknik clustering dan klasifikasi guna mengatasi kekurangan metode konvensional dan menghasilkan pengelompokkan yang lebih tepat sasaran. Kerangka ini menjadi pedoman dalam mengarahkan setiap tahap penelitian agar tujuan penelitian dapat dicapai secara sistematis dan terukur. Kerangka berfikir dapat dilihat pada *Gambar 2.1 Kerangka Berfikir*

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



Gambar 2. 1 Kerangka Berfikir

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)  
BAB III METODELOGI PENELITIAN



### 3.1. Analisa Sistem

#### 3.1.1. Analisa Sistem Yang Berjalan

Analisis terhadap sistem yang sedang berjalan bertujuan untuk memahami kondisi saat ini dalam pengelompokan kemampuan calistung siswa. Saat ini, sistem yang digunakan masih bergantung pada penilaian manual, yang sering kali tidak konsisten dan kurang objektif. Dengan menerapkan metode machine learning, diharapkan dapat meningkatkan efisiensi dan akurasi dalam mengelompokkan siswa berdasarkan tingkat kemampuan mereka.

Pengelompokan kemampuan calistung siswa saat ini menggunakan pendekatan tradisional, seperti penilaian manual dan observasi langsung oleh guru. Pendekatan ini menghadapi beberapa kendala, antara lain:

- 1) Subjektivitas. Penilaian yang bergantung pada persepsi guru dapat menyebabkan bias.
- 2) Efisiensi Waktu. Proses manual memakan waktu yang cukup lama, sehingga menyulitkan untuk segera mengidentifikasi kebutuhan siswa.
- 3) Data yang Terbatas. Data yang dikumpulkan sering kali tidak mencakup berbagai aspek keterampilan siswa, seperti membaca, menulis, dan berhitung secara terpadu.

Oleh karena itu, penting untuk mengembangkan sistem yang lebih efisien dan objektif guna menghasilkan pengelompokan yang lebih akurat dan relevan.

#### 3.1.2. Alternatif Pemecahan Masalah


Beberapa alternatif solusi yang diusulkan untuk mengatasi permasalahan ini meliputi:

- 1) Penggunaan Metode *Hybrid*, Mengadopsi metode *hybrid* dengan menggabungkan algoritma *K-Means* untuk pengelompokan siswa berdasarkan kemampuan, serta algoritma *K-Nearest Neighbors (KNN)* untuk klasifikasi. Pendekatan ini diharapkan dapat meningkatkan akurasi

## Protected by PDF Anti-Copy Free

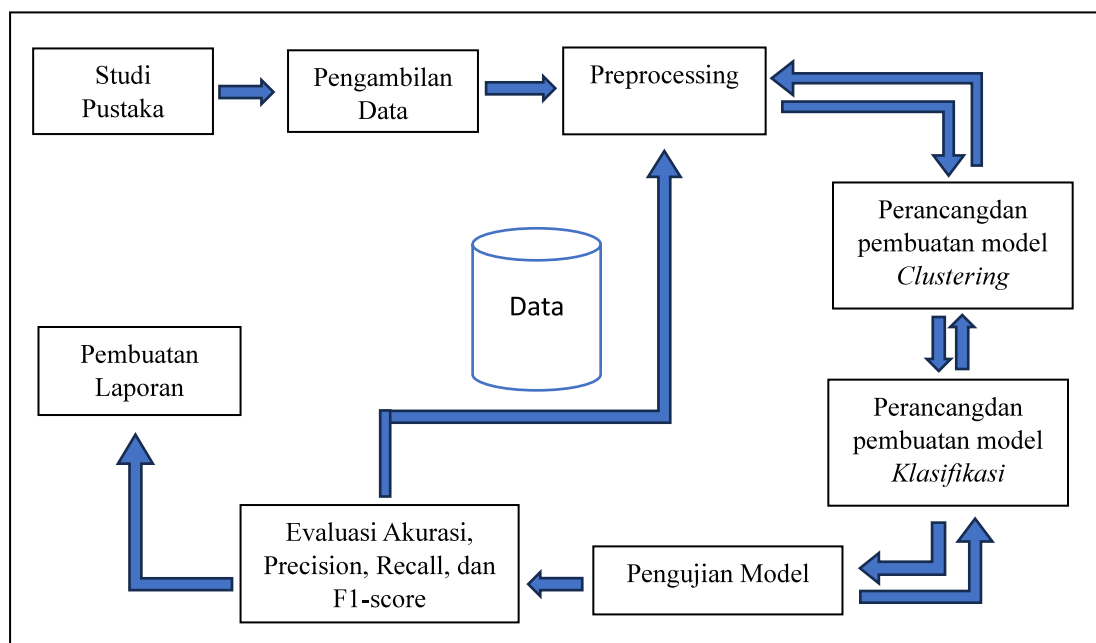
(Upgrade to Pro Version to Remove the Watermark)

dalam mengidentifikasi kelompok siswa sesuai dengan kemampuan mereka.

- 2) Implementasi Sistem  s Pemrogramman *Python*. Menerapkan sistem berbasis pemrogramman *Python* yang memudahkan guru dalam mengumpulkan, menganalisis, dan mengakses data siswa secara real-time. Sistem ini diharapkan dapat mempercepat proses pengambilan keputusan dan mendukung pengelolaan pembelajaran yang lebih efektif.
- 3) Pelatihan untuk Guru. Menyediakan program pelatihan bagi guru agar dapat memahami dan memanfaatkan hasil analisis data dalam merancang strategi pembelajaran. Pelatihan ini bertujuan membantu guru menyesuaikan metode pengajaran dengan kebutuhan spesifik setiap kelompok siswa berdasarkan hasil pengelompokan.

### 3.1.3. Metode Analisa

Analisis data dalam penelitian ini menggunakan pendekatan berbasis machine learning, dengan tujuan untuk mengelompokkan kemampuan siswa dalam membaca, menulis, dan berhitung (calistung) di Sekolah Dasar Kota Lubuklinggau. Tahapan analisis data meliputi gambar dibawah ini:



Gambar 3. 1 Analisis Konsep Penelitian

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Gambar 3.1 Analisis Konsep Penelitian menjelaskan alur proses analisis

data dalam penelitian yang menggunakan pendekatan berbasis machine learning. Berikut adalah tahapan penjelasan yang terlibat:

### 1) Studi Pustaka

Tahap awal di mana peneliti melakukan kajian literatur untuk memahami teori dan penelitian sebelumnya yang relevan dengan topik penelitian.

### 2) Pengumpulan Data

Mengumpulkan data yang diperlukan dari sumber-sumber yang relevan, seperti observasi, wawancara, dokumen dan Pustaka.

### 3) Preprocessing Data

Proses ini meliputi pembersihan dan pengolahan data agar siap digunakan dalam model machine learning. Ini termasuk menghapus data yang tidak relevan, menangani nilai yang hilang, dan normalisasi data menggunakan *min-max scaler*.

### 4) Pengembangan Model

Di tahap ini, peneliti memilih mengembangkan model machine learning yaitu dengan clustering dan klasifikasi, untuk menganalisis data.

### 5) Penerapan Pembelajaran Model

Model yang telah dikembangkan diterapkan pada data yang telah diproses untuk menghasilkan prediksi atau analisis lebih lanjut.

### 6) Evaluasi Akurasi

Setelah penerapan, model dievaluasi menggunakan metrik seperti akurasi, presisi, recall, dan F1-score untuk menilai kinerja model.

### 7) Penyusunan Laporan

Tahap akhir di mana peneliti menyusun laporan hasil penelitian, berdasarkan temuan rekomendasi dan analisis yang dilakukan.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 3.2. Teknik Pemilihan Informan (Populasi dan Sampel)

#### 3.2.3. Teknik Pengumpulan Data

Populasi dalam penelitian ini adalah siswa di dua Sekolah Dasar Kota Lubuklinggau yaitu SD Negeri 48 Lubuklinggau dan SD Negeri 23 Lubuklinggau pada tahun ajaran 2023/2024, dengan fokus pada siswa yang sedang dalam proses pengembangan kemampuan membaca, menulis, dan berhitung (calistung). Sampel Data yang digunakan dalam penelitian ini diambil dari hasil tes calistung yang telah dilakukan sebelumnya oleh siswa. Data yang dikumpulkan adalah nama siswa, jenis kelamin, asal sekolah, nilai membaca, nilai menulis, dan menghitung. Data yang didapat dari dua Sekolah Dasar Kota Lubuklinggau berjumlah 508 data siswa yang akan dijadikan sebagai sampel penelitian ini. Sampel data yang didapat terlihat pada *Table 3.1 Sampel Data Penelitian* dibawah ini :

Tabel 3. 1 Sampel Data Penelitian

No	Nama	P/L	Sekolah	Nilai 1	Nilai 2	Nilai 3
1	Alifa syakia	P	SD N 48	75	80	78
2	Atika putri	P	SD N 48	80	80	75
3	Kayra nasya	P	SD N 48	75	80	79
4	Boy candra	L	SD N 48	60	65	60
5	Cinta larasati	P	SD N 48	65	75	70
6	Dafa ananda	L	SD N 48	70	80	79
7	Dedek noelo	L	SD N 48	75	80	79
8	Delisa sabila	P	SD N 48	70	80	70
9	Denjaka halim	L	SD N 48	60	65	60
10	Diki saputra	L	SD N 48	80	80	70
11	Dimas nabil	L	SD N 48	70	65	60
12	Dio deandra	L	SD N 48	70	80	70
13	Diva Salsabila	P	SD N 48	80	80	75
14	Dorra inasya bela	P	SD N 48	80	80	76
15	Elvira syakilla	P	SD N 48	80	80	79
16	Adistiara	P	SD N 48	70	80	79
17	Yudha pranata	L	SD N 48	60	60	60
18	Frisnu adi	L	SD N 48	70	80	70
19	Gian novel	L	SD N 48	80	80	75
20	Iis saputri	P	SD N 48	80	80	76
21	Indri dahlia	P	SD N 48	80	80	79
22	Juwita intan	P	SD N 48	70	80	75

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

23	Karina latasya	P	SD N 48	65	65	60
24	Karisa isabella		SD N 48	80	80	79
25	Kavin trio		SD N 48	70	70	70
26	Aldi syaputra		SD N 48	75	80	79
27	Asyila zahrana		SD N 48	75	75	75
28	Anggi lisa	P	SD N 23	61	61	63
29	Aprilia	P	SD N 23	61	62	63
30	Chika meilan putri	P	SD N 23	72	67	70
31	Devri Simanjuntak	L	SD N 23	76	75	78
32	Glen fadly	L	SD N 23	69	68	70
33	Kevin Alfarizi	L	SD N 23	64	69	67
34	Naura Nabila	P	SD N 23	72	74	65
35	Nopia Adinda	P	SD N 23	65	67	64
36	Putri Tarissa	P	SD N 23	78	73	75
37	Tania	P	SD N 23	71	63	65
38	Ade Pratama	L	SD N 23	80	80	80
39	Andika Saputra	L	SD N 23	70	72	72
40	Bali Anugrah	L	SD N 23	70	70	70
41	Brayen Govala	L	SD N 23	70	70	70
42	Chika Aulia	P	SD N 23	86	87	85
43	Dani Erlanda	L	SD N 23	72	70	70
44	Jacko	L	SD N 23	70	72	70
45	Keyla Aprilian	P	SD N 23	79	76	78
46	M. Azka Jefrino	L	SD N 23	70	70	70
47	M. Firlo DwiAriyanto	L	SD N 23	75	73	72
48	Maxel Andika	L	SD N 23	70	72	68
49	Muhammad Marvel	L	SD N 23	80	78	79
50	Rangga Zahindra	L	SD N 23	70	72	70
51	Resti Marilina	P	SD N 23	80	84	84
52	Sofiah Anggela	P	SD N 23	72	70	70
53	Abdul Fakhi	L	SD N 23	83	83	82
54	Aqila Nadhira	P	SD N 23	82	82	81
55	Cristiandi Pratama	L	SD N 23	78	78	78
56	Jihan Sastra	L	SD N 23	81	81	80
57	Kelvin Frido	L	SD N 23	80	80	80
58	M. Sultan Raffasyah	L	SD N 23	74	74	71
59	Melati Puspa Sari	P	SD N 23	78	78	77
60	Muhammad Raditya	L	SD N 23	81	81	78
61	Redi Afriansyah	L	SD N 23	68	68	69
62	Rizki Pranata	L	SD N 23	80	80	78
63	Steven Kavisa	L	SD N 23	68	68	67
64	Pinkan Prisillia	P	SD N 23	78	78	78

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

65	Bunga Citra	P	SD N 23	77	77	78
66	Handika Saputra		SD N 23	73	73	72
67	Enji Herdiansyah		SD N 23	83	83	82
68	Aldo Permada		SD N 23	88	88	82
69	Anggun Sania		SD N 23	81	81	76
70	Annisa Azzahra	P	SD N 23	66	79	68
71	Azka Aris Pratama	L	SD N 23	85	85	86
72	Deko	L	SD N 23	82	83	82
73	Dhamar Kusuma	L	SD N 23	78	82	79
74	Dwinalirdhan	L	SD N 23	83	84	81
75	Farel Saputra	L	SD N 23	85	83	84
76	M. Zidan Al Fthir	L	SD N 23	81	82	81
77	Muhammad Aziz	L	SD N 23	75	80	80
78	Muhammad Gibran	L	SD N 23	82	83	78
79	Muhammad Riski	L	SD N 23	78	79	77
80	Nazella Aura	P	SD N 23	75	79	75
81	Nur Denta	L	SD N 23	81	79	81
82	Paris Lamaju	L	SD N 23	76	82	77
83	Syaka Farish	L	SD N 23	77	79	77
84	Zaki Islah Salam	L	SD N 23	89	81	88
85	Dimas Apriyansah	L	SD N 23	50	57	68
86	Aidil Samudra	L	SD N 23	65	63	66
87	Aiko Nasyilah	L	SD N 23	65	63	66
88	Aldi Prayoga	L	SD N 23	75	70	76
89	Andre Saputra	L	SD N 23	63	63	64
90	Arksha Virendha	L	SD N 23	63	63	64
91	Damara Atma	L	SD N 23	70	70	71
92	Desmi Aprilia	P	SD N 23	70	70	71
93	Dhea putri	P	SD N 23	80	77	81
94	Habibi muhammad	L	SD N 23	63	63	64
95	Khairulnissa	P	SD N 23	80	73	81
96	Khuza Al-fiqh	L	SD N 23	70	65	71
97	M. Johan	L	SD N 23	65	65	66
98	Mia Agusti	P	SD N 23	70	70	71
99	Muhammad dwi	L	SD N 23	63	63	64
100	Novita Lestari	P	SD N 23	63	63	64
101	Raffa Adytia	L	SD N 23	70	70	71
102	Shora inara	P	SD N 23	63	63	64
103	Shoza izareen	P	SD N 23	70	65	71
104	Siti Marwah	P	SD N 23	63	63	64
105	Vika Damayanti	P	SD N 23	70	70	71
106	Zaki Hamizan	L	SD N 23	80	80	81

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

107	Abdul Qodir	L	SD N 23	83	83	84
108	Andriyan		SD N 23	81	81	82
109	Anindita Khairin		SD N 23	81	70	82
110	Aska Aqila		SD N 23	83	82	84
111	Devi Putri		SD N 23	72	75	73
112	Jiban Kustaha	P	SD N 23	76	76	77
113	M. Duha abdila	L	SD N 23	80	78	81
114	M. Febriansyah	L	SD N 23	72	70	73
115	Merando	L	SD N 23	71	70	72
116	Muhammad Rizki	L	SD N 23	75	75	76
117	Muhammad Sendi	L	SD N 23	70	70	71
118	Nauval Alexi	L	SD N 23	81	82	82
119	Qur Rahman	L	SD N 23	70	75	71
120	Rania Humaria	P	SD N 23	82	83	83
121	Revatina Novariza	P	SD N 23	71	70	72
122	Sintia Anggraini	P	SD N 23	80	80	81
123	Syakila Waromah	P	SD N 23	81	82	82
124	Abdi Bastian	L	SD N 48	80	78	80
125	Abdurrahman	L	SD N 48	82	77	81
126	Abid alfiandra	L	SD N 48	79	75	80
127	Adelia Julianti	P	SD N 48	78	74	80
128	Adiel satriaman	L	SD N 48	77	80	82
129	Adinda Safira	P	SD N 48	80	79	82
130	Afnan Yudhoyono	P	SD N 48	81	77	82
131	Agustia Safitri	P	SD N 48	83	79	82
132	Adha Pian muza	L	SD N 48	75	78	80
133	Ahmad Hisyam Mursyid	L	SD N 48	77	79	81
134	Ahmad oktaviansyah	L	SD N 48	76	78	80
135	Aisyah	P	SD N 48	79	78	80
136	Alfarizki Dylan	L	SD N 48	80	79	82
137	Allea sagita	P	SD N 48	81	76	83
138	Alvarez Ramadhan	L	SD N 48	81	79	82
139	Amelia Putri	P	SD N 48	82	80	81
140	Amira hidayatun	P	SD N 48	83	81	82
141	Andini Cahfitri	P	SD N 48	79	78	80
142	Anisa serlina	P	SD N 48	75	79	81
143	Annisa wilanda	P	SD N 48	77	77	81
144	Anisha Anita	P	SD N 48	79	80	82
145	Apri Wulandari	P	SD N 48	80	77	81
146	Aqilla Wulandari	P	SD N 48	81	80	83
147	Arka Shaputra	L	SD N 48	80	80	83
148	Arya dinata Henry	L	SD N 48	82	80	82

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

149	Asyifa Septiana	P	SD N 48	80	79	81
150	Arka dena		SD N 48	75	78	80
151	Rizki prana		SD N 48	78	77	80
152	Muhammad Riduansyah		SD N 48	80	81	80
153	Nayla Aprilia		SD N 48	84	79	82
154	Nobel Aflah	L	SD N 48	85	79	80
155	Nur rafah	L	SD N 48	85	76	81
156	Pita Nurjanah	P	SD N 48	76	77	80
157	Puput Eka Putri	P	SD N 48	78	77	80
158	Putra mansa	L	SD N 48	77	77	80
159	Putri Aulia Sari	P	SD N 48	79	75	80
160	Rachelma Zaskiyah	P	SD N 48	80	77	82
161	Raihan Chairul	L	SD N 48	81	79	82
162	Rapa alfaginza	L	SD N 48	81	79	80
163	Raziq Hanan	L	SD N 48	82	79	83
164	Ridho Achmad	L	SD N 48	83	80	84
165	Safira Koldya	L	SD N 48	76	77	80
166	Sakhia Raisah	P	SD N 48	77	77	82
167	Shakila Fauziah	P	SD N 48	78	75	80
168	Sofia putri	P	SD N 48	79	77	80
169	Sultan Arif Dinata	L	SD N 48	77	75	82
170	Syarif Ilham	L	SD N 48	79	77	80
171	Syifa Nabila	P	SD N 48	76	77	80
172	Veбри Erina	P	SD N 48	79	78	84
173	Wisnu Arsyia	L	SD N 48	80	76	82
174	Zahira Alisa	P	SD N 48	81	75	80
175	Zaulan Elza	P	SD N 48	82	80	83
176	Zian Rizki	L	SD N 48	79	76	80
177	Ziya Azzahra	P	SD N 48	77	77	80
178	M ikhsan Bay	L	SD N 48	78	77	82
179	Keyla Putri Ramadhani	P	SD N 48	77	65	80
180	Kaysha Azizah	P	SD N 48	78	68	80
181	Kaysha Putri	P	SD N 48	75	65	79
182	Kenzo pandu Ariansyah	L	SD N 48	80	69	81
183	Khairina Cahyani	P	SD N 48	69	70	80
184	Lamondi antoso	L	SD N 48	80	66	79
185	Laoneis Daniyah	L	SD N 48	80	76	81
186	M hafiz ahsani	L	SD N 48	75	65	69
187	M. Nabihal	L	SD N 48	80	78	70
188	M. Varendra Hanif	L	SD N 48	75	65	80
189	M. Wildan abqori	L	SD N 48	78	70	80
190	Muhammad Ripki	L	SD N 48	78	70	79

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

191	Muhammad Rizki	L	SD N 48	80	77	78
192	Muhammad Zidan		SD N 48	76	76	79
193	Muhammad Azriel		SD N 48	76	79	80
194	Muhammad delki		SD N 48	81	78	80
195	Muhammad Fajri		SD N 48	77	80	81
196	Muhammad Haris	L	SD N 48	80	77	79
197	Muhammad kinnand	L	SD N 48	75	80	81
198	Muhammad Nurrohim	L	SD N 48	77	78	80
199	Muhammad Raditya	L	SD N 48	75	78	80
200	M. Vino Abdul Qodir	L	SD N 48	76	78	80
201	Muhammad Azka	L	SD N 48	79	77	80
202	Nabita Nur aini	P	SD N 48	80	78	81
203	Nasywa putriani	P	SD N 48	80	77	82
204	Nayshila Azkia Zahrana	P	SD N 48	80	83	84
205	Naira Adila Putri	P	SD N 48	77	81	79
206	Andi Prasetyo	L	SD N 48	85	78	90
207	Siti Aisyah	P	SD N 48	88	82	85
208	Budi Santoso	L	SD N 48	75	80	70
209	Rina Wulandari	P	SD N 48	90	85	88
210	Joko Susanto	L	SD N 48	70	75	80
211	Lina Marlina	P	SD N 48	92	90	95
212	Dika Ramadhan	L	SD N 48	80	77	82
213	Nia Kurniawati	P	SD N 48	78	80	76
214	Rudi Hartono	L	SD N 48	85	88	84
215	Ayu Lestari	P	SD N 48	89	90	91
216	Fajar Nugroho	L	SD N 48	76	74	78
217	Tika Sari	P	SD N 48	91	89	92
218	Eko Prabowo	L	SD N 48	82	80	81
219	Dinda Safitri	P	SD N 48	87	85	90
220	Rian Aditya	L	SD N 48	79	76	80
221	Sari Melati	P	SD N 48	93	91	94
222	Dimas Setiawan	L	SD N 48	88	84	87
223	Nisa Farah	P	SD N 48	85	82	89
224	Arif Hidayat	L	SD N 48	74	78	75
225	Putri Ayu	P	SD N 48	90	92	91
226	Bagus Pramono	L	SD N 48	81	79	83
227	Citra Lestari	P	SD N 48	84	88	86
228	Rudianto	L	SD N 48	77	75	78
229	Intan Permatasari	P	SD N 48	92	90	93
230	Denny Kurniawan	L	SD N 48	78	76	80
231	Siti Nurhaliza	P	SD N 48	89	87	90
232	Yudi Hartono	L	SD N 48	82	80	81

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

233	Rika Amalia	P	SD N 23	90	92	94
234	Agus Setyawan		SD N 23	76	74	77
235	Lila Sari		SD N 23	88	85	89
236	Farhan Rizky		SD N 23	75	73	76
237	Nanda Pratiwi		SD N 23	91	89	92
238	Iwan Setiawan	L	SD N 23	80	78	81
239	Diah Ayu	P	SD N 23	86	84	88
240	Rudi Prabowo	L	SD N 23	79	77	80
241	Sari Indah	P	SD N 23	92	91	93
242	Bayu Nugraha	L	SD N 23	83	80	82
243	Nia Kurnia	P	SD N 23	88	87	90
244	Fandi Ardiansyah	L	SD N 23	76	75	74
245	Cinta Melati	P	SD N 23	89	91	90
246	Dika Santoso	L	SD N 23	78	76	80
247	Rina Sari	P	SD N 23	90	88	91
248	Asep Hidayat	L	SD N 23	82	80	84
249	Tania Putri	P	SD N 23	85	87	90
250	Joni Prabuwono	L	SD N 23	74	75	73
251	Vina Lestari	P	SD N 23	91	89	92
252	Riko Adi	L	SD N 23	80	82	81
253	Nia Safira	P	SD N 23	88	86	89
254	Fajar Kurniawan	L	SD N 23	77	75	76
255	Cinta Rahayu	P	SD N 23	92	90	94
256	Arman Setiawan	L	SD N 23	81	79	80
257	Lila Melati	P	SD N 23	86	88	87
258	Denny Prabowo	L	SD N 23	78	76	77
259	Siti Aulia	P	SD N 23	90	91	93
260	Rudi Setiawan	L	SD N 23	79	78	80
261	Intan Permata	P	SD N 23	92	90	91
262	Bayu Pramudya	L	SD N 23	83	81	82
263	Nanda Citra	P	SD N 23	88	86	89
264	Fandi Nugroho	L	SD N 23	75	73	76
265	Diah Kurnia	P	SD N 23	89	91	90
266	Andi Saputra	L	SD N 23	80	78	81
267	Sari Indah	P	SD N 23	91	89	92
268	Rian Hidayat	L	SD N 23	76	75	78
269	Tika Melati	P	SD N 23	87	85	90
270	Eko Prabowo	L	SD N 23	82	80	81
271	Dinda Safitri	P	SD N 23	93	91	94
272	Farhan Rizki	L	SD N 23	79	76	80
273	Nia Kurnia	P	SD N 23	88	86	89
274	Rudianto	L	SD N 23	75	73	76

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

275	Cinta Rahayu	P	SD N 23	90	92	91
276	Dika Pramono		SD N 23	78	76	79
277	Vina Lestari		SD N 23	91	88	92
278	Joni Setiawan		SD N 23	82	80	81
279	Tania Permata		SD N 23	86	84	87
280	Arman Hidayat	L	SD N 23	79	77	80
281	Lila Sari	P	SD N 23	92	90	93
282	Fajar Kurniawan	L	SD N 23	83	81	82
283	Nanda Citra	P	SD N 23	88	86	89
284	Rudi Prabowo	L	SD N 23	76	75	74
285	Siti Nurhaliza	P	SD N 23	89	91	90
286	Dika Santoso	L	SD N 23	81	79	83
287	Rina Sari	P	SD N 23	84	88	86
288	Asep Hidayat	L	SD N 23	77	75	78
289	Tania Putri	P	SD N 23	92	90	93
290	Joni Prabuwono	L	SD N 23	78	76	80
291	Vina Lestari	P	SD N 23	89	87	90
292	Riko Adi	L	SD N 23	82	80	81
293	Nia Safira	P	SD N 23	90	92	94
294	Fajar Kurniawan	L	SD N 23	76	74	77
295	Cinta Rahayu	P	SD N 23	88	85	89
296	Arman Setiawan	L	SD N 23	75	73	76
297	Lila Melati	P	SD N 23	91	89	92
298	Denny Prabowo	L	SD N 23	80	78	81
299	Siti Aulia	P	SD N 23	86	84	88
300	Rudi Setiawan	L	SD N 23	79	77	80
301	Intan Permata	P	SD N 23	92	90	91
302	Bayu Pramudya	L	SD N 23	83	81	82
303	Nanda Citra	P	SD N 23	88	86	89
304	Fandi Nugroho	L	SD N 23	75	73	76
305	Diah Rosdiana	P	SD N 23	89	91	90
306	Fahri Setiawan	L	SD N 23	72	75	70
307	Gina Amelia	P	SD N 23	90	92	91
308	Hani Salsabila	P	SD N 23	78	76	79
309	Ika Rahayu	P	SD N 23	82	80	81
310	Juliansyah Prabowo	L	SD N 23	80	78	82
311	Kiki Marisa	P	SD N 23	87	85	88
312	Lia Anggraeni	P	SD N 23	75	74	76
313	Maya Sari	P	SD N 23	88	90	89
314	Nina Kurniawan	P	SD N 23	81	82	80
315	Omar Fikri	L	SD N 23	73	70	72
316	Putu Wira	L	SD N 23	85	87	84

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

317	Qori Anisa	P	SD N 23	90	91	92
318	Rina Amalia		SD N 23	76	75	74
319	Sari Rahma		SD N 23	89	88	87
320	Tina Yuliana		SD N 23	84	82	85
321	Udin Pramono		SD N 23	70	72	71
322	Vina Safira	P	SD N 23	80	78	82
323	Wira Purnama	L	SD N 23	78	75	76
324	Lila Puspita	P	SD N 23	84	83	85
325	Miko Prabowo	L	SD N 23	79	80	78
326	Nia Kurnia	P	SD N 23	92	91	90
327	Oka Ramadhan	L	SD N 23	76	74	75
328	Puspa Anggraini	P	SD N 23	87	88	86
329	Qais Alif	L	SD N 48	85	84	83
330	Rani Indah	P	SD N 48	90	92	91
331	Sandi Prabowo	L	SD N 48	78	76	77
332	Tika Melati	P	SD N 48	81	82	80
333	Ujang Setiawan	L	SD N 48	80	78	79
334	Vira Sari	P	SD N 48	86	84	85
335	Widiansyah Pramono	L	SD N 48	75	73	74
336	Xena Puspita	P	SD N 48	88	90	89
337	Yudha Prasetyo	L	SD N 48	84	82	83
338	Zaki Ramadhan	L	SD N 48	80	78	79
339	Alia Rahmawati	P	SD N 48	91	92	90
340	Budi Santoso	L	SD N 48	76	75	74
341	Chika Amalia	P	SD N 48	84	86	85
342	Danu Hidayat	L	SD N 48	79	80	78
343	Elina Safira	P	SD N 48	90	91	89
344	Farhan Akbar	L	SD N 48	75	74	73
345	Gina Amelia	P	SD N 48	85	84	86
346	Hendra Setiawan	L	SD N 48	88	89	87
347	Ika Sari	P	SD N 48	82	81	80
348	Joni Prabowo	L	SD N 48	78	76	77
349	Kiki Lestari	P	SD N 48	90	92	91
350	Luthfi Ramadhan	L	SD N 48	75	73	74
351	Melisa Indah	P	SD N 48	88	89	90
352	Niko Prabowo	L	SD N 48	82	81	80
353	Oki Saputra	L	SD N 48	87	85	86
354	Qila Nadhira	P	SD N 48	89	90	88
355	Rudi Setiawan	L	SD N 48	81	80	82
356	Siti Nurjanah	P	SD N 48	85	84	83
357	Taufik Hidayat	L	SD N 48	90	91	92
358	Uli Rahmawati	P	SD N 48	77	76	75

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

359	Vicky Prasetyo	L	SD N 48	81	80	82
360	Winda Lestari		SD N 48	84	86	85
361	Xandra Putri		SD N 48	89	90	88
362	Yudi Prabowo		SD N 48	78	75	77
363	Zainab Anisa		SD N 48	85	84	83
364	Aditia Nugraha	L	SD N 48	90	92	91
365	Bayu Wicaksono	L	SD N 48	76	78	77
366	Celine Dwi	P	SD N 48	82	80	81
367	Daren Prabowo	L	SD N 48	88	90	89
368	Elia Rahma	P	SD N 48	79	81	78
369	Fira Salsabila	P	SD N 48	85	84	86
370	Guntur Alamsyah	L	SD N 48	80	82	81
371	Hilda Mawar	P	SD N 48	90	88	89
372	Iwan Setiawan	L	SD N 48	75	77	76
373	Jannah Nabila	P	SD N 48	82	80	81
374	Kevin Aditya	L	SD N 48	88	86	90
375	Lila Puspita	P	SD N 48	84	83	85
376	Miko Prabowo	L	SD N 48	79	80	78
377	Nia Kurnia	P	SD N 48	92	91	90
378	Oka Ramadhan	L	SD N 48	76	74	75
379	Puspa Anggraini	P	SD N 48	87	88	86
380	Qais Alif	L	SD N 48	85	84	83
381	Rani Indah	P	SD N 48	90	92	91
382	Sandi Prabowo	L	SD N 48	78	76	77
383	Tika Melati	P	SD N 48	81	82	80
384	Ujang Setiawan	L	SD N 48	80	78	79
385	Vira Sari	P	SD N 48	86	84	70
386	Widiansyah Pramono	L	SD N 48	75	73	74
387	Xena Puspita	P	SD N 48	88	90	89
388	Yudha Prasetyo	L	SD N 48	84	82	83
389	Ariana Salsabila	P	SD N 48	85	78	92
390	Bintang Citra	P	SD N 48	92	85	90
391	Cahya Rini	P	SD N 48	80	77	88
392	Daffa Riyadi	L	SD N 48	70	65	75
393	Elang Pratama	L	SD N 48	88	80	85
394	Fira Zahra	P	SD N 48	95	90	93
395	Galang Arjuna	L	SD N 48	82	76	80
396	Hana Pratiwi	P	SD N 48	87	82	86
397	Iqbal Akbar	L	SD N 48	78	70	82
398	Jasmine Putri	P	SD N 48	85	88	91
399	Khalil Fikri	L	SD N 48	90	85	89
400	Larasati Kania	P	SD N 48	80	78	84

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

401	Muhammad Alfi	L	SD N 48	88	87	94
402	Nabila Sari		SD N 48	82	80	79
403	Omar Dwi		SD N 48	93	90	92
404	Puspa Dewi		SD N 48	75	70	73
405	Qorya Nabila		SD N 48	90	92	91
406	Rizki Danang	L	SD N 48	85	80	85
407	Sari Nurul	P	SD N 48	92	88	94
408	Taufik Rian	L	SD N 48	78	74	76
409	Umi Kalsum	P	SD N 48	90	86	89
410	Vira Oktaviani	P	SD N 48	82	75	80
411	Wira Prasetya	L	SD N 48	95	90	93
412	Xena Khairani	P	SD N 48	89	85	87
413	Yusuf Masykur	L	SD N 48	80	72	79
414	Zahra Fadilah	P	SD N 48	92	91	88
415	Aditya Gunawan	L	SD N 48	85	83	84
416	Bela Anjani	P	SD N 48	95	92	90
417	Cintya Dwi	P	SD N 48	88	80	85
418	David Angga	L	SD N 48	78	76	79
419	Elisabeth Wulandari	P	SD N 48	86	82	83
420	Fadil Ghifari	L	SD N 48	94	91	89
421	Gina Safira	P	SD N 48	85	88	90
422	Hendra Prabowo	L	SD N 48	88	85	92
423	Indra Setiawan	L	SD N 48	80	75	80
424	Jihan Alifa	P	SD N 48	92	89	87
425	Kiki Ramadhani	P	SD N 48	76	80	79
426	Lutfi Hanif	L	SD N 48	89	86	90
427	Maya Hidayati	P	SD N 48	84	88	89
428	Naufal Rizky	L	SD N 48	82	85	80
429	Olia Nabila	P	SD N 48	78	79	84
430	Putra Jaya	L	SD N 48	95	90	96
431	Qila Nurul	P	SD N 48	88	84	82
432	Rena Salwa	P	SD N 48	80	78	80
433	Syafiq Azhari	L	SD N 48	84	83	86
434	Tika Suryani	P	SD N 48	91	85	88
435	Uli Mawarni	P	SD N 48	78	82	79
436	Vino Pranata	L	SD N 48	93	90	92
437	Wulan Indah	P	SD N 48	92	88	85
438	Xander Fauzi	L	SD N 23	85	82	88
439	Yumna Zahra	P	SD N 23	94	91	90
440	Zaki Ramadhan	L	SD N 23	79	74	78
441	Alya Kamilia	P	SD N 23	83	80	86
442	Budi Santosa	L	SD N 23	90	89	91

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

443	Citra Dian	P	SD N 23	84	86	87
444	Dwi Saputra		SD N 23	87	83	85
445	Eka Novita		SD N 23	92	88	90
446	Fendy Marwan		SD N 23	85	83	84
447	Galih Nugraha		SD N 23	80	78	82
448	Heni Widyawati	P	SD N 23	92	91	90
449	Ilham Riza	L	SD N 23	88	85	89
450	Januar Tegar	L	SD N 23	85	80	88
451	Kartika Putri	P	SD N 23	80	77	81
452	Liana Sari	P	SD N 23	85	82	83
453	Mirza Iskandar	L	SD N 23	87	88	90
454	Nina Halimah	P	SD N 23	92	86	85
455	Oki Fauzi	L	SD N 23	91	84	86
456	Pia Maharani	P	SD N 23	80	79	82
457	Raka Aji	L	SD N 23	92	89	94
458	Sella Riza	P	SD N 23	86	85	82
459	Tino Dimas	L	SD N 23	84	80	87
460	Umi Wulandari	P	SD N 23	82	84	85
461	Vivi Citra	P	SD N 23	87	90	89
462	Wira Andika	L	SD N 23	93	86	92
463	Xenia Puspita	P	SD N 23	91	85	88
464	Yara Hania	P	SD N 23	85	82	84
465	Zainal Abidin	L	SD N 23	79	74	78
466	Amira Salma	P	SD N 23	95	89	91
467	Bima Arya	L	SD N 23	90	83	89
468	Cita Arum	P	SD N 23	91	84	85
469	Dian Agustin	P	SD N 23	86	80	87
470	Erik Rahman	L	SD N 23	90	85	87
471	Farida Hariani	P	SD N 23	80	77	81
472	Gusti Agung	L	SD N 23	89	88	90
473	Hendri Saputra	L	SD N 23	87	86	91
474	Indah Sari	P	SD N 23	84	81	82
475	Juliana Zuraida	P	SD N 23	93	89	88
476	Kurniawan Dimas	L	SD N 23	79	75	78
477	Luthfi Rinaldi	L	SD N 23	88	82	80
478	Mira Fadila	P	SD N 23	92	89	90
479	Nisa Nuraini	P	SD N 23	87	84	85
480	Oscar Tio	L	SD N 23	93	90	91
481	Putri Anggun	P	SD N 23	88	85	86
482	Rifki Hadi	L	SD N 23	79	77	81
483	Sinta Nurbaya	P	SD N 23	91	86	88
484	Tatang Farhan	L	SD N 23	82	79	83

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

485	Ulya Siti	P	SD N 23	86	84	90
486	Vina Rizki		SD N 23	89	86	85
487	Wandy Aditya		SD N 23	85	80	82
488	Yuliana Candra		SD N 23	93	89	87
489	Akarshi Kumar		SD N 23	85	78	92
490	Nanan Hajamudin	L	SD N 23	90	85	88
491	Nabila Rahman	P	SD N 23	80	87	91
492	Mei Linda	P	SD N 23	95	90	93
493	Xavin Axendra	L	SD N 23	70	76	81
494	Ana Subrata	P	SD N 23	88	92	79
495	Idris Omer	L	SD N 23	75	81	86
496	Zara Elzani	P	SD N 23	82	87	78
497	Kaitan hamildan	L	SD N 23	89	84	90
498	Siti Amina	P	SD N 23	91	89	87
499	Leila Hadjani	P	SD N 23	93	85	83
500	Tommas Yunan	L	SD N 23	78	80	74
501	Yulia Ivanova	P	SD N 23	95	91	88
502	Hadiyah Al-Farouq	P	SD N 23	76	70	95
503	Enriq Diaz	L	SD N 23	80	77	82
504	Aidan Canfaiz	L	SD N 23	85	81	86
505	Farhan Aziz	L	SD N 23	89	94	90
506	Sofia Villanueva	P	SD N 23	86	79	91
507	Kamilah Abdul	P	SD N 23	90	82	87
508	Jalya doni	L	SD N 23	84	88	82

### 3.2.3. Perancangan Instrumen Quisioner

#### 1) Observasi

Penulis memperoleh data dengan melakukan observasi langsung ke dua Sekolah Dasar dan melihat gambaran nyata mengenai kesulitan yang dialami siswa dalam menguasai kemampuan dasar calistung tersebut.

#### 2) Wawancara (interview)

Peneliti melakukan wawancara langsung dengan kepala sekolah dan guru di dua Sekolah Dasar Kota Lubuklinggau yang mengajar mata pelajaran calistung, untuk mendapatkan informasi mendalam mengenai kemampuan membaca, menulis, dan berhitung siswa, dan ditemukan bahwa siswa sekolah dasar kelas 1 sampai dengan kelas 6 kemampuan calistungnya masih rendah.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 3) Metode Dokumentasi

Pengambilan data dilakukan dengan cara mendokumentasikan dokumen dan data terkait.

### 4) Metode Pustaka

Data yang penulis dapat berasal dari berbagai sumber yang ada pada suatu buku, laporan, jurnal dan literatur yang berhubungan dengan permasalahan yang diangkat.

## 3.2.3. Teknik Analisa

### 3.2.3.1. Pengolahan Data

Pengolahan data dalam penelitian ini melibatkan serangkaian langkah mulai dari preprocessing data hingga penerapan algoritma *K-Means* dan metode *hybrid* yang digabungkan dengan algoritma lain seperti *K-Nearest Neighbors (KNN)*.

#### 1) Preprocessing Data

Sebelum algoritma diterapkan, data perlu melalui tahapan preprocessing untuk membersihkan dan mempersiapkan data. Langkah-langkahnya meliputi:

- a) Normalisasi Data: Menstandarkan data agar berada dalam rentang yang sama, menggunakan *Min-Max Scaling*:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3)$$

- b) Penanganan Missing Data: Jika ada data yang hilang, maka data tersebut dapat diimputasi atau diabaikan tergantung pada jumlah data yang hilang.
- c) Transformasi Fitur: Jika data memiliki outlier, transformasi seperti log transformation dapat diterapkan.

#### 2) Clustering dengan K-Means

Metode *K-Means* ini digunakan untuk mengelompokkan siswa berdasarkan kemampuan calistung mereka. Algoritma *K-Means* akan membagi siswa ke dalam beberapa kelompok berdasarkan kesamaan dalam data kemampuan, sehingga memudahkan guru dalam

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

mengidentifikasi kelompok siswa yang memerlukan pendekatan pembelajaran yang

### 3) Klasifikasi dengan PDF

Setelah perolehan data, dapat dilakukan klasifikasi untuk menempatkan siswa variabel dalam cluster yang sesuai. Penelitian ini menggunakan metode *K-Nearest Neighbors (KNN)* untuk mengklasifikasikan siswa baru berdasarkan data yang telah dikelompokkan. Teknik ini membantu dalam menentukan kategori kemampuan siswa berdasarkan jarak terdekat dengan data yang ada.

#### 3.2.3.2. Pengujian Model

Pengujian model ini dilakukan untuk memastikan keakuratan hasil pengelompokan dan klasifikasi. Pengujian dilakukan melalui beberapa pendekatan untuk mendapatkan performa terbaik dari metode *hybrid* yang digunakan.

##### 1) Metode Elbow (*Elbow Method*)

Metode *Elbow* digunakan untuk pengujian model pada *Clustering K-Means* untuk menentukan jumlah cluster optimal yang sesuai dengan dataset. Metode ini didasarkan pada nilai *Sum of Squared Errors (SSE)* atau *inertia* yang menggambarkan seberapa baik setiap titik data dalam cluster cocok dengan centroidnya. Langkah-langkah dalam Metode *Elbow*:

- a) Jalankan algoritma *K-Means* dengan berbagai jumlah cluster  $k$ .
- b) Hitung SSE untuk setiap  $k$ , dengan rumus:

$$SSE = \sum_{i=1}^n \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (4)$$

- c) di mana  $x$  adalah data dalam cluster  $C_i$  dan  $\mu_i$  adalah centroid cluster tersebut.
- d) Plot grafik antara  $k$  dan *SSE*.
- e) Tentukan titik "*elbow*" (titik di mana perubahan *SSE* mulai melambat), yang menandakan jumlah cluster optimal.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)  
2) Confusion Matrix

Setelah pengevaluasian digunakan algoritma klasifikasi *KNN* maka pengujian menggunakan *Confusion Matrix* digunakan untuk menilai akurasi prediksi kelas. *Confusion Matrix* menampilkan jumlah prediksi yang benar dan salah untuk setiap kelas yang diklasifikasikan.

*Confusion Matrix* terdiri dari:

- True Positive (TP)** : Prediksi benar untuk kelas positif.
- False Positive (FP)** : Prediksi salah untuk kelas positif.
- True Negative (TN)** : Prediksi benar untuk kelas negatif.
- False Negative (FN)**: Prediksi salah untuk kelas negatif.

Metrik yang dapat dihitung dari *Confusion Matrix* adalah:

a) **Akurasi:**

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

b) **Presisi:**

$$Presisi = \frac{TP}{TP+FP} \quad (6)$$

c) **Recall:**

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

d) **F1-Score**

$$F1 - Score = 2 \cdot \frac{Precision+Recall}{Precision \cdot Recall} \quad (8)$$

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 3.3. Tempat dan Waktu Penelitian

#### 3.3.1. Tempat Penelitian

Penelitian ini dilaksanakan bertempat di dua Sekolah Dasar Kota Lubuklinggau, yaitu SD Negeri 8 Kota Lubuklinggau dan SD Negeri 23 Kota Lubuklinggau.

#### 3.3.2. Waktu Penelitian

Perkiraan pelaksanaan penelitian ini dijadwalkan untuk dimulai pada bulan September 2024, sebagaimana terinci dalam *Tabel 3.2 Waktu Penelitian*. Rencana ini mencakup jadwal kerja yang telah disusun untuk proyek penelitian yang dimulai pada tanggal tersebut.

Tabel 3. 2 Waktu Penelitian

No	KEGIATAN	Waktu Penelitian																			
		September 2024				Oktober 2024				November 2024				Desember 2024				Januari 2025			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Analisis Permasalahan	■	■																		
2	Pengajuan Judul		■	■																	
3	Penbuatan Proposal & Bimbingan				■	■	■	■	■	■	■	■	■								
4	Ujian Proposal & Perbaikan												■	■	■						
5	Pembuatan Sistem															■	■				
6	Skripsi																■	■			
7	Ujian Akhir																			■	■

#### 3.3.3. Alat dan Bahan

##### 1) Alat

Dalam penelitian ini menggunakan alat-alat sangat mendukung untuk suksesnya penelitian ini yang berupa *Hardware* dan *software* sebagai berikut :

- a. Perangkat Keras (*Hardware*)
  - a) Laptop Asus
  - b) Printer Canon Pixma MG2570S
- b. Perangkat Lunak (*Software*)
  - a) Windows 10 Profesional 64 Bit
  - b) Microsoft Office 2019

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

c) *B.python*

d) *Mendeley Destop*

e) *Jupyter Notebo*



### 2) Bahan

Dalam penelitian ini menggunakan Bahan-Bahan yang sangat mendukung penelitian ini yakni :

- a. Tinta Printer (*Black dan Colour*)
- b. Kertas A4 80 Gram

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### BAB IV HASIL PENELITIAN DAN PEMBAHASAN



#### 4.1 Gambaran Umum

##### 4.1.1. Gambaran Umum (Tempat Penelitian)

Penelitian ini bertujuan untuk mengatasi kesulitan dalam pengelompokan kemampuan membaca, menulis, dan berhitung (calistung) siswa di tingkat sekolah dasar, dengan fokus pada SD Negeri 48 dan SD Negeri 23 di Kota Lubuklinggau. Kedua sekolah ini menghadapi tantangan dalam mengelompokkan siswa secara manual, yang sering kali kurang efisien, membutuhkan waktu dan tenaga yang signifikan, serta menghasilkan tingkat akurasi yang rendah.

Untuk meningkatkan efisiensi dan keakuratan dalam proses tersebut, penelitian ini menerapkan metode hybrid berbasis machine learning. Metode ini mengintegrasikan algoritma K-Means Clustering untuk membagi siswa ke dalam beberapa kelompok berdasarkan kemampuan mereka, dan algoritma K-Nearest Neighbors (KNN) untuk memprediksi kelompok siswa baru berdasarkan data yang telah ada. Dengan pendekatan ini, sistem mampu melakukan pengelompokan data siswa yang ada serta memberikan prediksi yang akurat bagi siswa baru.

Hasil penelitian diharapkan dapat menjadi solusi praktis dalam mendukung proses pembelajaran di kedua sekolah tersebut dengan menyediakan alat berbasis data yang lebih efektif. Sistem ini dirancang untuk membantu guru merancang strategi pengajaran yang lebih relevan dengan kebutuhan masing-masing kelompok siswa, sehingga dapat meningkatkan efektivitas pengajaran dan hasil belajar secara keseluruhan. Penelitian ini menggunakan data nilai tes membaca, menulis, dan berhitung siswa dari tahun ajaran 2023/2024 sebagai bahan analisis untuk menghasilkan pengelompokan yang lebih presisi.

Secara keseluruhan, penelitian ini berkontribusi dalam meningkatkan kualitas evaluasi pembelajaran di sekolah dasar dengan memanfaatkan teknologi modern. Pendekatan ini diharapkan mampu mengoptimalkan hasil

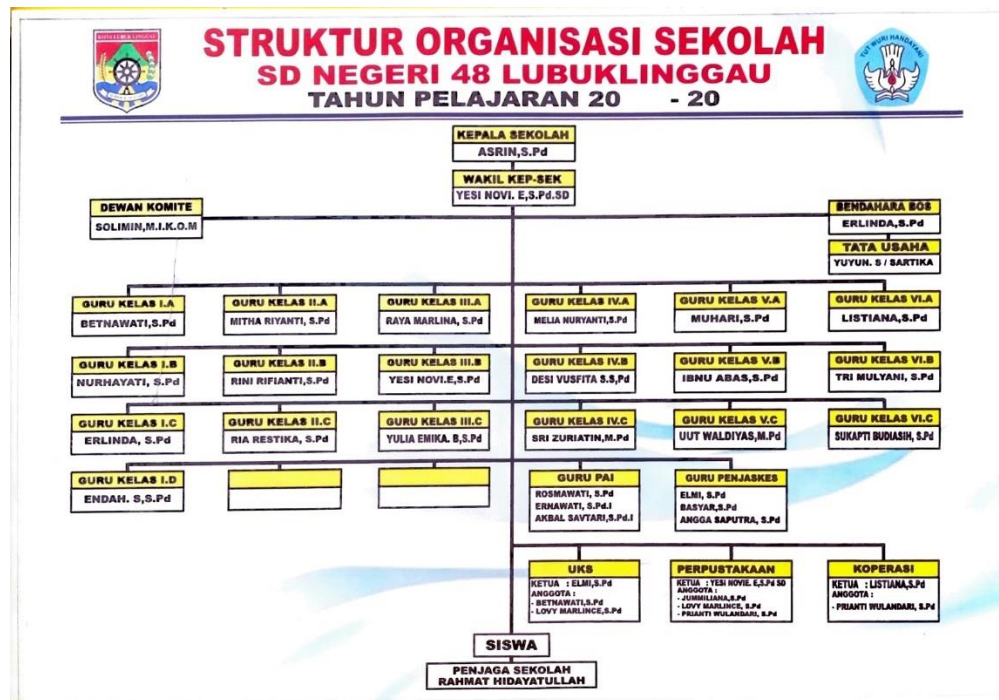
## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

belajar siswa secara individu maupun kelompok serta mendukung pengembangan kurikulum yang lebih adaptif terhadap kebutuhan siswa. Selain itu, penelitian ini membuka peluang bagi penelitian lanjutan tentang penerapan machine learning di bidang pendidikan untuk meningkatkan mutu pendidikan di Indonesia.

### 4.1.2. Struktur Organisasi

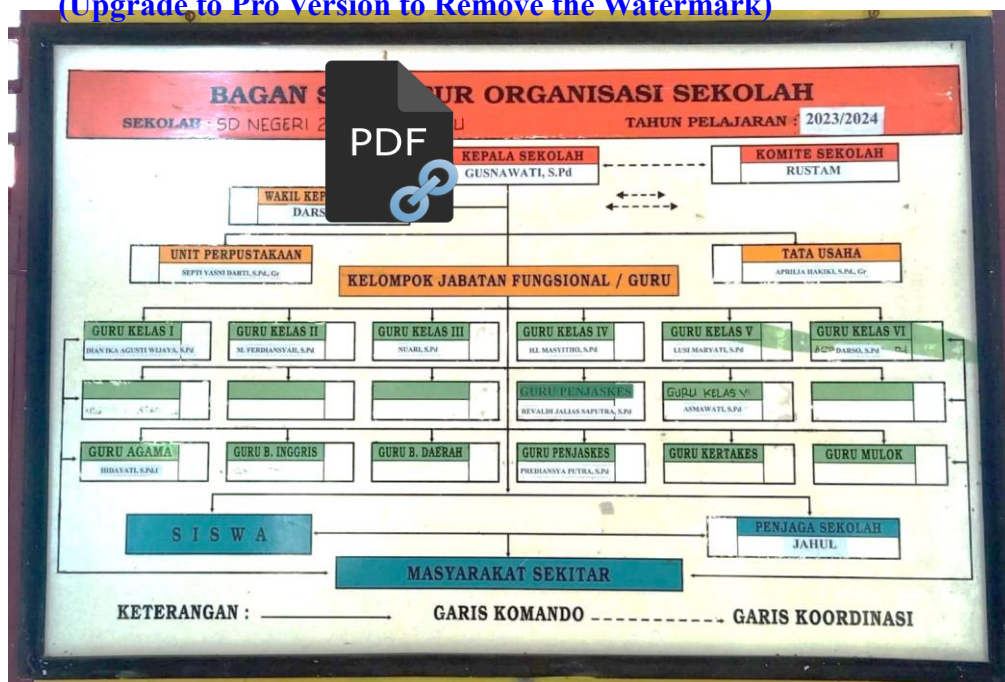
Struktur organisasi tempat penelitian ini adalah ada dua tempat penelitian sekolah dasar dimana Struktur organisasi sekolah dasar berperan penting dalam mendukung kelancaran operasional dan kegiatan pendidikan berjalan dengan efektif, termasuk kegiatan pembelajaran, administrasi, dan pengembangan siswa. Struktur Organisasi sekolah dasar bisa dilihat pada *Gambar 4.1 Struktur Organisasi Sekolah Dasar Negeri 48 Kota Lubuklinggau* dan *4.2 Struktur Organisasi Sekolah Dasar Negeri 23 Kota Lubuklinggau*



Gambar 4.1 Struktur Organisasi Sekolah Dasar Negeri 48 Kota  
Lubuklinggau

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



Gambar 4. 2 Struktur Organisasi Sekolah Dasar Negeri 23 Kota Lubuklinggau

### 4.2 Hasil Penelitian

Dalam penelitian ini, data kemampuan membaca, menulis, dan berhitung (calistung) siswa yang diperoleh dari SD Negeri 48 dan SD Negeri 23 di Kota Lubuklinggau berjumlah 508 data, dan untuk data pengujian berjumlah 200 sampel data telah dianalisis dengan menggunakan *metode hybrid* yang berbasis *machine learning*. Metode yang diterapkan mencakup algoritma *K-Means Clustering* untuk mengelompokkan siswa dan algoritma *K-Nearest Neighbors (KNN)* untuk melakukan klasifikasi.

Penelitian ini bertujuan untuk menganalisis kinerja model klustering atau pengelompokkan menggunakan K-Means pada dataset nilai siswa dan klasifikasi K-Nearest Neighbors (KNN) untuk pengujian model cluster pada data baru yang berjumlah 200 data sampel.

#### 4.2.1. Hasil Clustering *K-Means*

Pada hasil clustering, terbentuk 3 cluster yaitu (0, 1, dan 2). Hasil clustering yang terbentuk dari proses clustering ini dianalisis dan dikelompokkan yaitu mendapatkan hasil berikut:

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

- Cluster 0: Siswa dengan kemampuan calistung Menengah.
- Cluster 1: Siswa dengan kemampuan calistung Tinggi.
- Cluster 2: Siswa dengan kemampuan calistung rendah.

Jumlah siswa pada masing-masing kelompok berdasarkan cluster yang didapat yaitu:

- Nilai Menengah 267 siswa
- Nilai Tinggi 177 siswa
- Nilai Rendah 64 siswa

Analisis hasil centeroid dan nilai rata rata yang didapat adalah:

- Centeroid cluster 0

Nilai 1: 79.01

Nilai 2: 77.86

Nilai 3: 79.62

Ini adalah titik pusat untuk cluster yang berisi siswa dengan kinerja baik atau menengah. Dengan nilai rata-rata di atas 75.

- Centeroid cluster 1

Nilai 1: 88.70

Nilai 2: 86.98

Nilai 3: 88.66

Ini adalah titik pusat untuk cluster yang berisi siswa berprestasi tinggi. Dengan nilai rata-rata di atas 85.

- Centeroid cluster 2

Nilai 1: 68.22

Nilai 2: 68.62

Nilai 3: 68.58

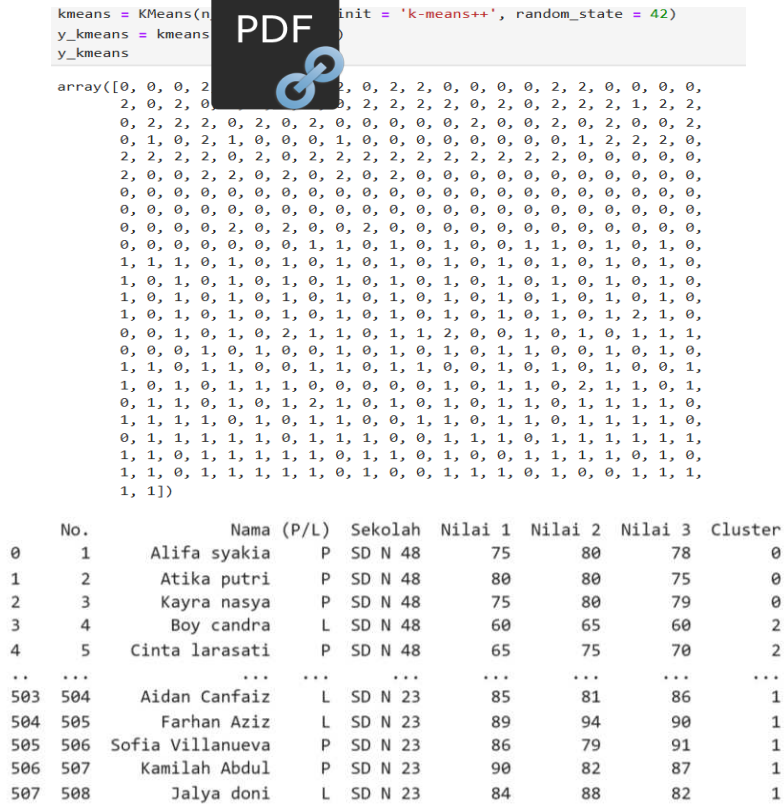
Ini adalah titik pusat untuk cluster yang berisi siswa dengan kinerja di bawah rata-rata. Siswa dalam cluster ini menunjukkan kinerja di bawah rata-rata, dengan nilai rata-rata di bawah 70. Ini menunjukkan bahwa mereka mungkin memerlukan perhatian lebih dalam proses pembelajaran.

Kriteria pengelompokan dilakukan menggunakan centroid (pusat klaster). Nilai siswa dibandingkan dengan centroid masing-masing

**Protected by PDF Anti-Copy Free**

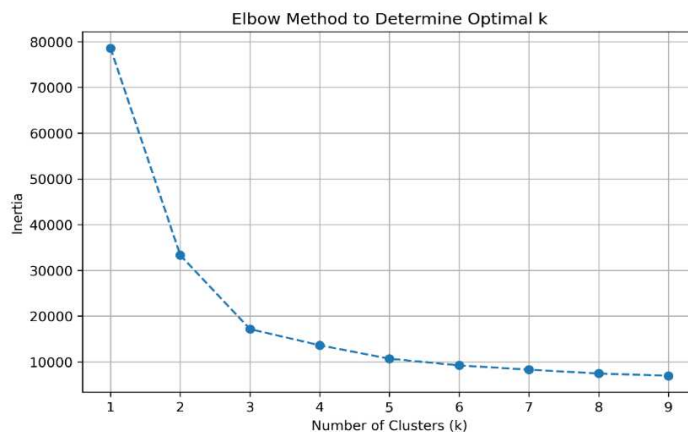
**(Upgrade to Pro Version to Remove the Watermark)**

klaster, dan siswa dimasukkan ke klaster yang paling dekat. Hasil dapat dilihat pada gambar selanjutnya berikut:



Gambar 4. 3 Hasil Data Cluster Dengan K-Means

Pada pengujian *cluster optimal* dengan menggunakan *Elbow Method* didapatkan *cluster Optimal=3*. Dapat dilihat pada gambar 4.3 *Grafik Metode Elbow Menentukan Nilai K Optimal*. Dimana titik grafik mulai melambat (menyerupai siku) menunjukkan jumlah cluster yang optimal.

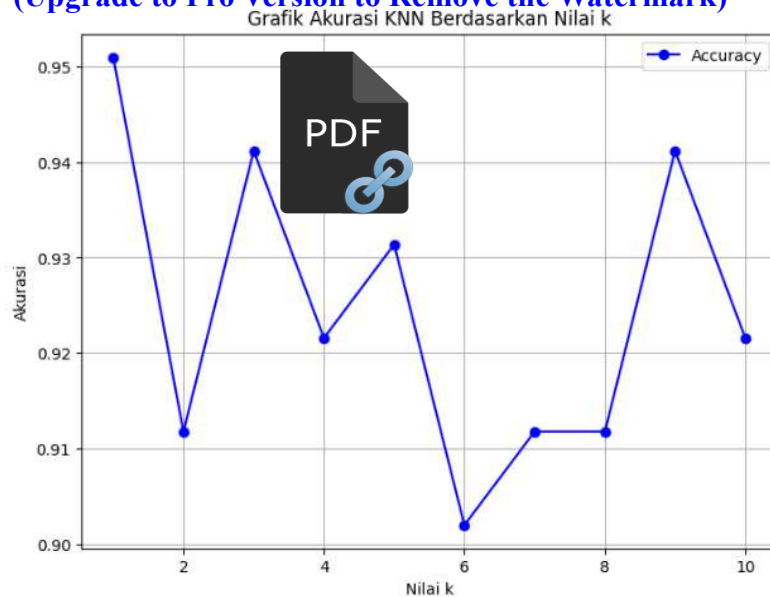


Gambar 4. 4 Grafik Metode Elbow Menentukan Nilai K Optimal



## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

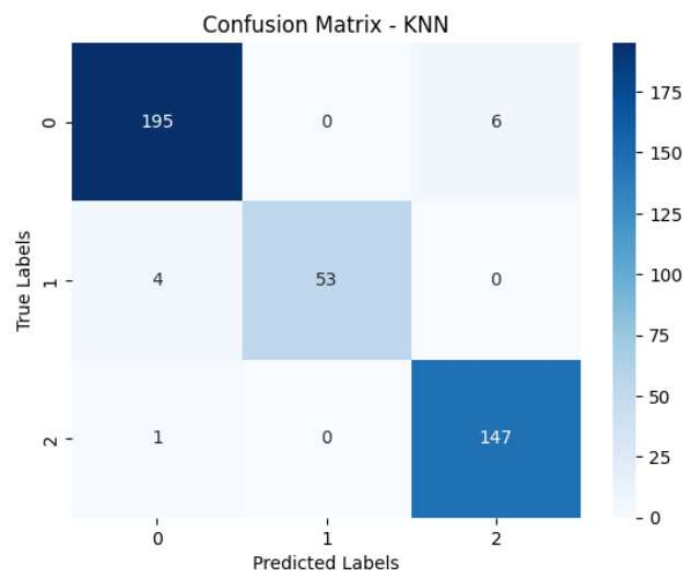


```
# Mencetak nilai k dengan akurasi terbaik
best_k = k_values[np.argmax(accuracies)]
print(f"Nilai k terbaik adalah: {best_k} dengan akurasi: {max(accuracies)}")
```

Nilai k terbaik adalah: 1 dengan akurasi 0.9509803921568627

Gambar 4. 6 Grafik Akurasi K-Nearest Neighbors (KNN) Nilai K Terbaik

Pada gambar grafik ini terlihat bahwa Nilai k terbaik yang ditemukan adalah 1, dengan akurasi tertinggi sebesar 95.09%. Selanjutnya dilakukan uji Confusion matrix menunjukkan bahwa model memiliki performa yang baik dalam mengklasifikasikan data, dengan precision dan recall yang tinggi untuk setiap kelas. Dapat dilihat dari gambar berikut:



Gambar 4. 7 Grafik Confusion Matrix

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Hasil pengujian model Confusion Matrik perhitungan manual accuracy, precision, recall, dan f1-score:

- a) **True Positive (TP)** : Prediksi benar untuk kelas positif.
- b) **False Positive (FP)** : Prediksi salah untuk kelas positif.
- c) **True Negative (TN)** : Prediksi benar untuk kelas negatif.
- d) **False Negative (FN)** : Prediksi salah untuk kelas negatif.

Metrik yang dapat dihitung dari Confusion Matrix adalah:

Tabel 4. 1 Confusion Matrix

True Labels\Predicted Labels	0	1	2
0	195	0	6
1	4	53	0
2	1	0	147

a) **Accuracy:**

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

Rumus ini memang lebih umum digunakan untuk kasus **binary classification** di mana ada **TP (True Positive)**, **TN (True Negative)**, **FP (False Positive)**, dan **FN (False Negative)**. Namun, pada **multi-class classification** (seperti dalam gambar hasil confusion matrix), kita menghitung akurasi **secara keseluruhan**.

$$Akurasi = \frac{JumlahTotalPrediksiBenar(TP)}{Total\ Data} = \frac{395}{406} = 0.97 \text{ (97\%)}$$

- **TP** = Jumlah prediksi benar untuk semua kelas → nilai diagonal:

TP untuk kelas 0 = 195

TP untuk kelas 1 = 53

TP untuk kelas 2 = 147

$$\text{Total TP} = 195+53+147 = 395$$

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

- **Total Data** = Semua elemen dalam matriks:

$$\text{Total Data} = 194 + 53 + 0 + 1 + 0 + 147 = 406$$



- b) **Precision:**

$$\text{Precision} = \frac{TP}{TP+FP} \quad (6)$$

### Kelas 0:

- TP = 195 (benar diprediksi kelas 0)
- FP = 4+1 = 5 (data kelas 1 dan 2 yang salah diprediksi sebagai kelas 0)

$$\text{Precision}_0 = \frac{195}{195 + 5} = \frac{195}{200} = 0.97$$

### Kelas 1:

- TP = 53
- FP = 0+0 = 0 (data kelas 0 dan 2 yang salah diprediksi sebagai kelas 1)

$$\text{Precision}_1 = \frac{53}{53 + 0} = \frac{53}{53} = 1.0$$

### Kelas 2:

- TP = 147
- FP = 6 + 0 = 6 (data kelas 0 dan 1 yang salah diprediksi sebagai kelas 2)

$$\text{Precision}_2 = \frac{147}{147 + 6} = \frac{147}{153} = 0.96$$

- c) **Recall:**

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

### Kelas 0:

- TP = 195
- FN = 0+6 = 6 (data kelas 0 yang salah diprediksi sebagai kelas 1 dan 2)

$$\text{Recall}_0 = \frac{195}{195 + 6} = \frac{195}{201} = 0.97$$

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

**Kelas 1:**

- TP = 53
- FN = 4 + 0 = 4 (data kelas 1 salah diprediksi sebagai kelas 0 dan 2)

$$Recall_1 = \frac{53}{53 + 4} = \frac{53}{57} = 0.93$$

**Kelas 2:**

- TP = 145
- FN = 1 + 0 = 1 (data kelas 2 yang salah diprediksi sebagai kelas 0 dan 1)

$$Recall_2 = \frac{145}{145 + 1} = \frac{145}{146} = 0.99$$

**d) F1-Score**

$$F1 - Score = 2 \cdot \frac{Precision + Recall}{Precision \cdot Recall} \quad (8)$$

**Kelas 0:**

$$F1 - Score = 2 \cdot \frac{0.97 + 0.97}{0.97 \cdot 0.97} = 2 \cdot \frac{0.9409}{1.94} = 0.97$$

**Kelas 1:**

$$F1 - Score = 2 \cdot \frac{1.0 + 0.93}{1.0 \cdot 0.93} = 2 \cdot \frac{0.93}{1.93} = 0.96$$

**Kelas 2:**

$$F1 - Score = 2 \cdot \frac{0.96 + 0.99}{0.96 \cdot 0.99} = 2 \cdot \frac{0.9504}{1.95} = 0.98$$

Hasil ini konsisten dengan *classification report* sebelumnya. Model memiliki performa tinggi dan bekerja dengan baik di setiap kelas.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

	precision	recall	f1-score	support
0		0.97	0.97	201
1		0.93	0.96	57
2		0.99	0.98	148
accuracy			0.97	406
macro avg	0.98	0.96	0.97	406
weighted avg	0.97	0.97	0.97	406

Gambar 4. 8 Hasil Akurasi, Precision, Recall dan F1-Score

Hasil evaluasi menunjukkan bahwa :

- **Accuracy:**

Akurasi menunjukkan persentase prediksi benar terhadap keseluruhan data.  $Accuracy = 0.97$  (97%).

- **Macro Avg:**

Rata-rata metrik *precision*, *recall*, dan *F1-Score* untuk semua kelas tanpa memperhatikan jumlah data dalam setiap kelas.  $Precision, Recall, F1-Score = 0.97$ .

- **Weighted Avg:**

Rata-rata metrik *precision*, *recall*, dan *F1-Score* dengan mempertimbangkan jumlah data dalam setiap kelas.  $Precision, Recall, F1-Score = 0.97$ .

Penelitian ini menunjukkan bahwa model KNN memberikan akurasi yang sangat baik dengan **akurasi 97%** dalam klasifikasi data siswa. Selain itu, *clustering* menggunakan *K-Means* berhasil mengelompokkan siswa berdasarkan nilai mereka, yang dapat digunakan untuk analisis lebih lanjut dalam pendidikan.

Hasil penelitian ini dapat digunakan sebagai dasar untuk pengembangan lebih lanjut dalam analisis data pendidikan dan penerapan teknik machine learning dalam konteks yang lebih luas.

Setelah itu dibuat inputan untuk data terbaru dengan nilai rata rata dan nilai calistung siswa. Dengan output sesuai cluster dan nilai siswa akan di kelompokkan dan di prediksi. Dapat dilihat dari gambar berikut:

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Masukkan nama Anda: Amanda Salsabila  
Masukkan nilai 1: 90  
Masukkan nilai 2: 100  
Masukkan nilai 3: 88  
Nama: Amanda Salsabila  
Nilai rata-rata: 92.67  
Kategori cluster berdasarkan dataset: Tinggi

Gambar 4. 9 Hasil Inputan Akhir

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.3 Pembahasan

#### 4.1.3. Penerapan Metode Analisis dan Validasi data

##### 4.1.3.1. Mengimpor *Library Python*

Mengimpor *library Python* bahasa pemrograman *Python* memungkinkan kita untuk memanfaatkan fungsi-fungsi yang ada dalam *library* tersebut. Tampilan pengimporan *library Python* dapat dilihat pada gambar berikut:

```
# Import semua Library yang diperlukan untuk KMeans dan KNN
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import random
import tensorflow as tf
from mpl_toolkits.mplot3d import Axes3D

# Preprocessing dan pemodelan
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.impute import SimpleImputer

# KMeans
from sklearn.cluster import KMeans

# KNN
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import (
    confusion_matrix, accuracy_score, classification_report,
    recall_score, precision_score
)
from sklearn.utils import shuffle

# Konfigurasi visualisasi
%matplotlib inline
```

Gambar 4. 10 Import Library

##### 4.1.3.2. Memuat data `read_csv`

Perintah `read_csv` dalam bahasa pemrograman *Python* digunakan untuk mengakses dan membaca file dataset yang berformat CSV (Comma Separated Values). Penggunaan `read_csv` dapat dilihat pada gambar berikut:

```
data = pd.read_csv('nilai_calistung.csv')
print(data.head())
```

	No.	Nama (P/L)	Sekolah	Nilai 1	Nilai 2	Nilai 3
0	1	Alifa syakia	P SD N 48	75	80	78
1	2	Atika putri	P SD N 48	80	80	75
2	3	Kayra nasya	P SD N 48	75	80	79
3	4	Boy candra	L SD N 48	60	65	60
4	5	Cinta larasati	P SD N 48	65	75	70

Gambar 4. 11 Read\_CSV

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.1.3.3. Statistik Deskriptif

Memberikan ringkasan statistik deskriptif, seperti rata-rata, deviasi standar, nilai minimum, maksimum, dan kuartil untuk kolom-kolom numerik dalam dataset. Informasi deskripsi pada data dapat dilihat dari gambar berikut:

```
data.describe()
```

	No.	Nilai 1	Nilai 2	Nilai 3
<b>count</b>	508.000000	508.000000	508.000000	508.000000
<b>mean</b>	254.500000	81.023622	79.872047	81.375984
<b>std</b>	146.791235	7.363179	7.015717	7.178918
<b>min</b>	1.000000	50.000000	57.000000	60.000000
<b>25%</b>	127.750000	77.000000	76.000000	78.000000
<b>50%</b>	254.500000	80.000000	80.000000	81.000000
<b>75%</b>	381.250000	87.000000	85.000000	87.000000
<b>max</b>	508.000000	95.000000	94.000000	96.000000

Gambar 4. 12 Data Describe

### 4.1.3.4. Informasi Dataset

Menampilkan informasi tentang dataset seperti jumlah baris, tipe data untuk setiap kolom, dan apakah terdapat data yang hilang (*missing values*). Informasi tentang dataset dapat dilihat pada gambar berikut:

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 508 entries, 0 to 507
Data columns (total 7 columns):
#   Column      Non-Null Count  Dtype
---  -
0   No.         508 non-null   int64
1   Nama        508 non-null   object
2   (P/L)      508 non-null   object
3   Sekolah    508 non-null   object
4   Nilai 1    508 non-null   int64
5   Nilai 2    508 non-null   int64
6   Nilai 3    508 non-null   int64
dtypes: int64(4), object(3)
memory usage: 27.9+ KB
```

Gambar 4. 13 Informasi Dataset

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.1.3.5. Membagi data menjadi Fitur dan Target

Membagi data menjadi fitur dan target dilakukan untuk mempersiapkan dataset untuk melakukan analisis atau pelatihan model. Dalam konteks ini, fitur adalah atribut yang digunakan untuk memprediksi, sedangkan target adalah nilai yang ingin diprediksi. Berikut adalah contoh cara membagi data menjadi fitur dan target dalam Python:

```
# Membagi data menjadi data training dan testing
X = data[['Nilai 1', 'Nilai 2', 'Nilai 3']]
y = data['Nama'] # Misalkan 'Nama' adalah Label (target)
```

Gambar 4. 14 Membagi Data Fitur dan Target

- X adalah subset dari dataset yang berisi fitur (atribut) yaitu kolom Nilai 1, Nilai 2, dan Nilai 3,
- sementara y adalah target atau label, misalnya kolom "Nama".

### 4.1.3.6. Membagi Data Training dan Testing

Membagi dataset menjadi 80% data latih (X\_train, y\_train) dan 20% data uji (X\_test, y\_test). Dan random\_state= 42 memastikan pembagian data selalu konsisten setiap kali kode dijalankan. Perintah pembagian data training dan testing dapat dilihat pada gambar berikut:

```
# bagi data training dan testing
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Mencetak ukuran data pelatihan
print("Ukuran data pelatihan:", x_train.shape)
print("Ukuran data pengujian:", y_train.shape)

Ukuran data pelatihan: (406, 8)
Ukuran data pengujian: (406,)
```

Gambar 4. 15 Pembagian Data Training dan Testing

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.1.4. Pengujian Hasil Analisa

#### 4.1.4.1. Clustering dengan K-

Clustering dilakukan untuk mengelompokkan siswa berdasarkan kemiripan pola calistung mereka. Berikut adalah langkah-langkah dan hasilnya:

#### 4.3.2.1.1. Hasil Clustering

Menginisialisasi model *K-Means* untuk membagi data menjadi 3 cluster, melatih model tersebut pada data X, dan kemudian mengembalikan label cluster untuk setiap titik data. Hasilnya disimpan dalam `y_kmeans`, yang dapat digunakan untuk analisis lebih lanjut atau visualisasi.

```
kmeans = KMeans(n_clusters= 3, init = 'k-means++', random_state = 42)
y_kmeans = kmeans.fit_predict(X)
y_kmeans
array([[0, 0, 0, 2, 2, 0, 0, 2, 2, 0, 2, 2, 0, 0, 0, 0, 2, 2, 0, 0, 0, 0,
        2, 0, 2, 0, 0, 2, 2, 2, 0, 2, 2, 2, 2, 0, 2, 0, 2, 2, 2, 1, 2, 2,
        0, 2, 2, 2, 0, 2, 0, 2, 0, 0, 0, 0, 0, 2, 0, 0, 2, 0, 2, 0, 0, 2,
        0, 1, 0, 2, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 2, 2, 2, 0,
        2, 2, 2, 2, 0, 2, 0, 2, 2, 2, 2, 2, 2, 2, 2, 2, 0, 0, 0, 0, 0,
        2, 0, 0, 2, 2, 0, 2, 0, 2, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 0, 0, 2, 0, 2, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
        0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0,
        1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0,
        1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1,
        1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1,
        0, 0, 1, 0, 1, 0, 2, 1, 1, 0, 1, 1, 2, 0, 0, 1, 0, 1, 0, 1, 1, 1,
        0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0,
        1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1,
        1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 1, 1, 0, 2, 1, 1, 0, 1,
        0, 1, 1, 0, 1, 0, 1, 2, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0,
        1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0,
        0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1,
        1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 0, 1, 0,
        1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1,
        1, 1])
```

Gambar 4. 16 Hasil Clustering

Pada gambar diatas ini merupakan hasil klasterisasi data menggunakan algoritma K-Means. Nilai-nilai pada array menunjukkan klaster mana data tersebut dimasukkan.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Pada kasus ini, terdapat 3 kluster yaitu (0, 1, dan 2). Hasil clustering yang terbelah dari proses clustering data menggunakan k-means. Berikut adalah hasilnya:

- Cluster 0: Siswa dengan kemampuan calistung Menengah.
- Cluster 1: Siswa dengan kemampuan calistung Tinggi.
- Cluster 2: Siswa dengan kemampuan calistung rendah.

### 4.3.2.1.2. Clustering dengan K Optimal

Melakukan clustering pada data siswa berdasarkan nilai mereka dalam tiga penilaian. Hasil clustering disimpan dalam DataFrame dan ditampilkan, sehingga kita dapat melihat nama siswa, nilai mereka, dan cluster yang ditugaskan kepada mereka. Perintah ini dapat dilihat pada gambar berikut:

```
# Load data
data = pd.read_csv('nilai_calistung.csv')
optimal_clusters = 3

# Lakukan clustering dengan k optimal
kmeans = KMeans(n_clusters=optimal_clusters, random_state=42)
data['Cluster'] = kmeans.fit_predict(X)
data[['Nama', 'Nilai 1', 'Nilai 2', 'Nilai 3', 'Cluster']]
```

	Nama	Nilai 1	Nilai 2	Nilai 3	Cluster
0	Alifa syakia	75	80	78	0
1	Atika putri	80	80	75	0
2	Kayra nasya	75	80	79	0
3	Boy candra	60	65	60	2
4	Cinta larasati	65	75	70	2
...	...	...	...	...	...
503	Aidan Canfaiz	85	81	86	1
504	Farhan Aziz	89	94	90	1
505	Sofia Villanueva	86	79	91	1
506	Kamilah Abdul	90	82	87	1
507	Jalya doni	84	88	82	1

508 rows × 5 columns

Gambar 4. 17 Clustering dengan K Optimal

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.3.2.1.3. Menghitung Rata Rata dan Menentukan Centeroid

Menentukan titik pusat atau nilai yaitu centeroid yang di temukan oleh k-means memberikan cara untuk mencari rata-rata dari setiap cluster. Dengan mengetahui posisi centroid, dapat digunakan untuk mencari rata-rata kinerja siswa dalam setiap kelompok. Penerapan ini dapat dilihat pada gambar berikut:

```
# Mengelompokkan berdasarkan cluster dan menghitung rata-rata untuk Membaca, Menulis, Menghitung
cluster_summary = data.groupby('cluster')[['Nilai 1', 'Nilai 2', 'Nilai 3']].mean()
print(cluster_summary)
```

	Nilai 1	Nilai 2	Nilai 3
Cluster			
0	79.007547	77.860377	79.622642
1	88.702247	86.977528	88.657303
2	68.215385	68.615385	68.584615

```
# Menampilkan centroid yang ditemukan setelah iterasi
print("Centroid yang ditemukan oleh K-Means:\n", kmeans.cluster_centers_)
```

```
Centroid yang ditemukan oleh K-Means:
[[79.00754717 77.86037736 79.62264151]
 [88.70224719 86.97752809 88.65730337]
 [68.21538462 68.61538462 68.58461538]]
```

Gambar 4. 18 Rata Rata dan Menentukan Centeroid

Analisis hasil centeroid dan nilai rata rata yang didapat adalah:

- Centeroid cluster 0

Nilai 1: 79.01

Nilai 2: 77.86

Nilai 3: 79.62

Ini adalah titik pusat untuk cluster yang berisi siswa dengan kinerja baik atau menengah. Dengan nilai rata-rata di atas 75.

- Centeroid cluster 1

Nilai 1: 88.70

Nilai 2: 86.98

Nilai 3: 88.66

Ini adalah titik pusat untuk cluster yang berisi siswa berprestasi tinggi. Dengan nilai rata-rata di atas 85.

- Centeroid cluster 2

Nilai 1: 68.22

Nilai 2: 68.62

Nilai 3: 68.58

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Ini adalah titik pusat untuk cluster yang berisi siswa dengan kinerja di bawah rata-rata. Siswa dalam cluster ini menunjukkan kinerja di bawah rata-rata, dengan nilai rata-rata di bawah 70. Ini menunjukkan bahwa mereka mungkin memerlukan perhatian lebih dalam proses pembelajaran.

### 4.3.2.1.4. Menentukan Centroid Akhir

Setelah proses clustering selesai, langkah selanjutnya adalah menentukan centroid akhir dari setiap cluster. Centroid ini merepresentasikan titik pusat dari setiap kelompok data, yang dihitung sebagai rata-rata dari semua titik data yang termasuk dalam cluster tersebut.

Setelah mendapatkan centroid akhir, kita dapat menganalisis dan menyimpan data terbaru untuk diklasifikasi dan melihat bagaimana data tersebut berinteraksi dengan cluster yang telah ditentukan. Dapat dilihat pada gambar berikut:

```
# Centroid akhir
print("Centroid Final:\n", kmeans.cluster_centers_)

# Data dengan klaster
print(data)
```

Centroid Final:

```
[[79.00754717 77.86037736 79.62264151]
 [88.70224719 86.97752809 88.65730337]
 [68.21538462 68.61538462 68.58461538]]
```

No.	Nama (P/L)	Sekolah	Nilai 1	Nilai 2	Nilai 3	Cluster
0	1 Alifa syakia	P SD N 48	75	80	78	0
1	2 Atika putri	P SD N 48	80	80	75	0
2	3 Kayra nasya	P SD N 48	75	80	79	0
3	4 Boy candra	L SD N 48	60	65	60	2
4	5 Cinta larasati	P SD N 48	65	75	70	2
..	...	...	...	...	...	...
503	504 Aidan Canfaiz	L SD N 23	85	81	86	1
504	505 Farhan Aziz	L SD N 23	89	94	90	1
505	506 Sofia Villanueva	P SD N 23	86	79	91	1
506	507 Kamilah Abdul	P SD N 23	90	82	87	1
507	508 Jalya doni	L SD N 23	84	88	82	1

Gambar 4. 19 Centroid Akhir

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.3.2.1.5. Menentukan Jumlah Optimal Cluster Menggunakan Metode Elbow

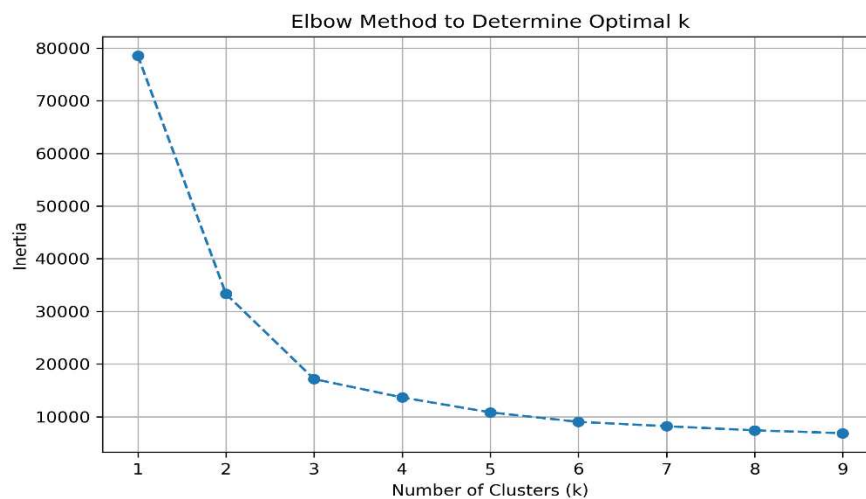
Metode *Elbow* digunakan untuk menentukan jumlah cluster optimal dengan mengplot nilai *inertia* (jarak kuadrat rata-rata dari titik data ke pusat cluster). Titik di mana grafik mulai melambat (menyerupai siku) menunjukkan jumlah cluster yang optimal. Perintah tersebut dapat dilihat pada gambar berikut:

```
# Select relevant numeric features for clustering
features = data[['Nilai 1', 'Nilai 2', 'Nilai 3']]

# Determine the optimal number of clusters using the elbow method
inertia = []
k_range = range(1, 10)

for k in k_range:
    kmeans = KMeans(n_clusters=k, random_state=42)
    kmeans.fit(features)
    inertia.append(kmeans.inertia_)

# Plot the elbow curve
plt.figure(figsize=(8, 5))
plt.plot(k_range, inertia, marker='o', linestyle='--')
plt.xlabel('Number of Clusters (k)')
plt.ylabel('Inertia')
plt.title('Elbow Method to Determine Optimal k')
plt.grid(True)
plt.show()
```



Gambar 4. 20 Jumlah Optimal Cluster Menggunakan Metode Elbow

Jumlah cluster optimal ditentukan menggunakan metode *Elbow*. Grafik *Within-Cluster-Sum-of-Squares* (WCSS) menunjukkan bahwa jumlah cluster optimal adalah  $k = 3$ .

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.3.2.1.6. Visualisasi 3D Hasil Clustering

Visualisasi 3D hasil clustering memberikan gambaran yang jelas tentang bagaimana data terdistribusi dalam ruang tiga dimensi. Dengan menggunakan Python dan matplotlib. Hasil visualisasi ditunjukkan pada gambar berikut:

```
# Inisialisasi dan Latih KMeans dengan 3 Kluster
kmeans = KMeans(n_clusters=3, random_state=42)
data = pd.DataFrame({'Cluster': kmeans.fit_predict(features)})

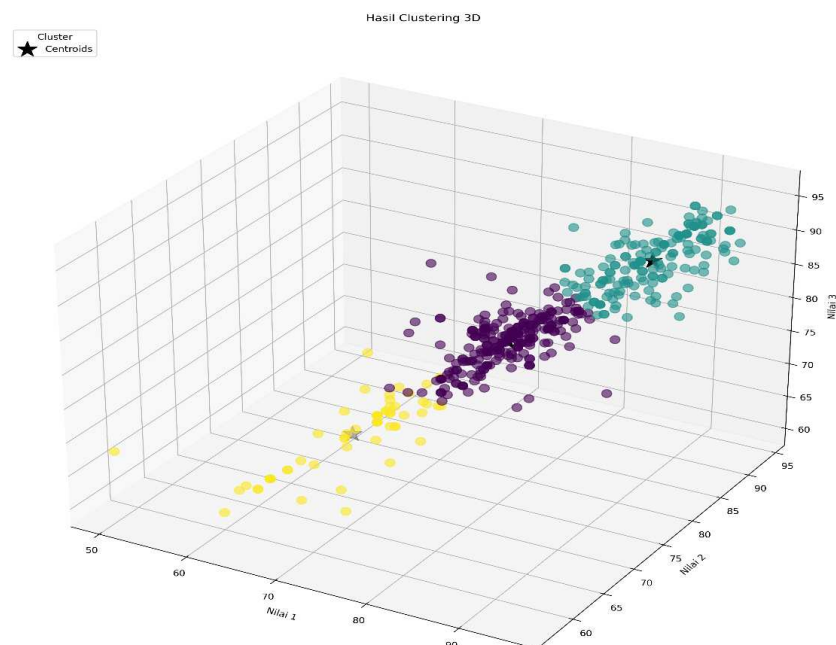
# Buat plot 3D
fig = plt.figure(figsize=(15, 15))
ax = fig.add_subplot(111, projection='3d')

# Scatter plot untuk Kluster
scatter = ax.scatter(
    features['Nilai 1'],
    features['Nilai 2'],
    features['Nilai 3'],
    c=data['Cluster'],
    cmap='viridis',
    s=100,
    alpha=0.6)

# Tambahkan centroid ke plot
centers = kmeans.cluster_centers_
ax.scatter(
    centers[:, 0], centers[:, 1], centers[:, 2],
    c='black', s=400, marker='*', label='Centroids')

# Atur Label dan judul
ax.set_title('Hasil Clustering 3D')
ax.set_xlabel('Nilai 1')
ax.set_ylabel('Nilai 2')
ax.set_zlabel('Nilai 3')
ax.legend(title='Cluster', loc='upper left')

# Tampilkan plot
plt.show()
```



Gambar 4. 21 Visualisasi 3D Hasil Clustering

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.3.2.1.7. Proses Mengakses data terbaru csv dan Penggabungan file

Proses membaca data csv baru yang sudah memiliki label agar dapat dilakukan klasifikasi dengan menggunakan algoritma knn dan sering kali dilakukan penggabungan file agar lebih efisien dan mudah digunakan untuk pengolahan data yang lebih kompleks. Dapat dilihat pada gambar berikut:

```
#Proses membaca data csv
terbaru=pd.read_csv('clustered_data_knn.csv', delimiter=',')
```

```
#Proses Penggabungan file dataset
df=pd.concat([terbaru],axis=0)
df=shuffle(df)
data = pd.DataFrame(df)
data
```

No.	Nama	(P/L)	Sekolah	Nilai 1	Nilai 2	Nilai 3	Cluster
79	80	Nazella Aura	P SD N 23	75	79	75	0
316	317	Qori Anisa	P SD N 23	90	91	92	1
485	486	Vina Rizki	P SD N 23	89	86	85	1
396	397	Iqbal Akbar	L SD N 48	78	70	82	0
167	168	Sofia putri	P SD N 48	79	77	80	0
...	...	...	...	...	...	...	...
106	107	Abdul Qodir	L SD N 23	83	83	84	0
270	271	Dinda Safitri	P SD N 23	93	91	94	1
348	349	Kiki Lestari	P SD N 48	90	92	91	1
435	436	Vino Pranata	L SD N 48	93	90	92	1
102	103	Shoza izareen	P SD N 23	70	65	71	2

Gambar 4. 22 Mengakses data terbaru csv dan Penggabungan file

### 4.3.2.1.8. Membuat Label Baru dari Cluster

Dengan memberikan label baru pada cluster, dapat lebih mudah memahami dan menjelaskan karakteristik dari setiap kelompok. Yaitu menengah(0), tinggi(1) dan rendah(2).

```
# buat label baru dari cluster
def label_cluster(cluster):
    if cluster == 0:
        return 'menengah'
    elif cluster == 1:
        return 'tinggi'
    elif cluster == 2:
        return 'rendah'
    else:
        return 'unknown'

data['D.Cluster'] = data['Cluster'].apply(label_cluster)
data
```

No.	Nama	(P/L)	Sekolah	Nilai 1	Nilai 2	Nilai 3	Cluster	D.Cluster
79	80	Nazella Aura	P SD N 23	75	79	75	0	menengah
316	317	Qori Anisa	P SD N 23	90	91	92	1	tinggi
485	486	Vina Rizki	P SD N 23	89	86	85	1	tinggi
396	397	Iqbal Akbar	L SD N 48	78	70	82	0	menengah
167	168	Sofia putri	P SD N 48	79	77	80	0	menengah
...	...	...	...	...	...	...	...	...
106	107	Abdul Qodir	L SD N 23	83	83	84	0	menengah
270	271	Dinda Safitri	P SD N 23	93	91	94	1	tinggi
348	349	Kiki Lestari	P SD N 48	90	92	91	1	tinggi
435	436	Vino Pranata	L SD N 48	93	90	92	1	tinggi
102	103	Shoza izareen	P SD N 23	70	65	71	2	rendah

Gambar 4. 23 Label Baru dari Cluster

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.3.2.1.9. Penghapusan Kolom yang Bernilai String

Proses penghapusan kolom yang bernilai string dan pembagian data menjadi fitur dan target penting dalam persiapan data untuk analisis data pembelajaran mesin. Dengan menghapus kolom yang tidak relevan dan memisahkan fitur dari target, dapat memastikan bahwa model dapat belajar dari data dengan cara yang efektif dan tidak terpengaruh oleh informasi yang tidak berguna. Dapat dilihat pada gambar berikut:

```
#Proses Penghapusan kolom yang bernilai string
#Membagi data menjadi fitur (X) dan target (y)
X=data.drop(columns=['Nama', 'D.Cluster'])
y=data['D.Cluster']
```

Gambar 4. 24 Penghapusan Kolom yang Bernilai String

### 4.3.2.1.10. Pembagian Data dan inialisasi Encode

Persiapan data untuk model pembelajaran mesin. Pembagian data menjadi dataset pelatihan 80% dan pengujian 20% serta encoding data kategorikal adalah langkah-langkah yang diperlukan untuk memastikan bahwa model dapat belajar dari data dengan cara yang efektif.

- (406, 8): Ini menunjukkan bahwa `x_train` memiliki 406 sampel dan 8 fitur.
- (102, 8): Ini menunjukkan bahwa `x_test` memiliki 102 sampel dan 8 fitur.
- (406,) : Ini menunjukkan bahwa `y_train` memiliki 406 label yang sesuai dengan data pelatihan.
- (102, ) : Ini menunjukkan bahwa `y_test` memiliki 102 label yang sesuai dengan data pengujian.

Dapat dilihat pada gambar berikut:

## Protected by PDF Anti-Copy Free

### (Upgrade to Pro Version to Remove the Watermark)

```

#Proses pembagian data training dan testing
x_train, x_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 42)

print(x_train.shape,x_test.shape,y_train.shape,y_test.shape)

# Inisialisasi encoder
label_encoder = LabelEncoder

# Lakukan encoding untuk data kategorikal
x_train = x_train.apply(lambda col: label_encoder.fit_transform(col) if col.dtypes == 'object' else col)
y_train = label_encoder.fit_transform(y_train)
x_test = x_test.apply(lambda col: label_encoder.fit_transform(col) if col.dtypes == 'object' else col)
y_test = label_encoder.fit_transform(y_test)

# Mencetak ukuran data pelatihan
print("Ukuran data pelatihan:", X_train.shape)
print("Ukuran data pengujian:", y_train.shape)

(406, 8) (102, 8) (406,) (102,)
Ukuran data pelatihan: (406, 3)
Ukuran data pengujian: (406,)

```

Gambar 4. 25 Pembagian Data dan inisialisasi Encode

#### 4.3.2.1.11. Klasifikasi Menggunakan *K-Nearest Neighbors (KNN)*

Setelah dilakukan *clustering k-means* didapatkan label Data hasil kluster tersebut dapat dijadikan data latih untuk *K-Nearest Neighbors (KNN)*. Setelah terbentuk model Selanjutnya *K-Nearest Neighbors (KNN)* digunakan untuk memprediksi kelas siswa dengan data sampel baru berjumlah 200 data. Untuk data training menggunakan dataset dengan data yang sudah dikelompokkan dan mempunyai label atau kelas. Dengan cara ini, *K-Nearest Neighbors (KNN)* dapat memberikan prediksi klasifikasi kemampuan *calistung* yang lebih akurat.

#### 4.3.2.1.12. Model *K-Nearest Neighbors (KNN)*

Proses pembuatan dan pelatihan model *K-Nearest Neighbors (KNN)*. Dapat dilihat pada gambar berikut:

```

#Membuat model KNN
AB= KNeighborsClassifier(n_neighbors=3)
AB.fit(x_train,y_train)

```

KNeighborsClassifier

KNeighborsClassifier(n\_neighbors=3)

Gambar 4. 26 Model *K-Nearest Neighbors (KNN)*

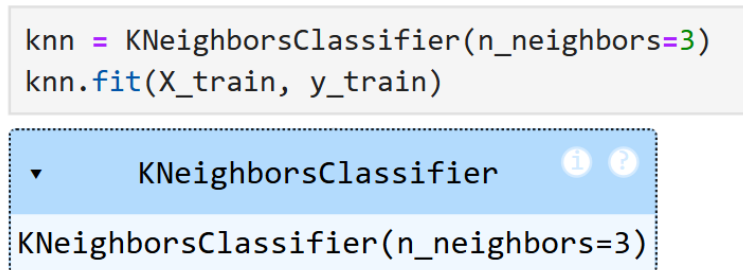
## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.3.2.1.13. Model *K-Nearest Neighbors (KNN) Prediksi Cluster*

Proses pembelajaran dan pelatihan model *K-Nearest Neighbors (KNN)* pada pengujian prediksi cluster pada data baru. Dapat dilihat pada gambar berikut:

```
knn = KNeighborsClassifier(n_neighbors=3)
knn.fit(X_train, y_train)
```



```
KNeighborsClassifier(n_neighbors=3)
```

Gambar 4. 27 Model (*KNN*) Prediksi Cluster Untuk Data Baru

### 4.3.2.1.14. Read CSV Data Sampel Baru Untuk Pengujian

Proses membaca data csv baru yaitu data sampel untuk pengujian model prediksi cluster pada data baru dapat dilihat pada gambar berikut:

```
# Data baru untuk pengujian
new_data200 = pd.read_csv('data_sampel_200.csv')
print(new_data200)
```

No.		Nama (P/L)	Sekolah	Nilai 1	Nilai 2	Nilai 3
0	1	Adi Handayani	P SD N 48	60	60	81
1	2	Fajar Haryono	P SD N 48	83	80	83
2	3	Eka Haryono	L SD N 48	72	100	80
3	4	Nana Saputra	L SD N 48	54	55	98
4	5	Eka Anggara	P SD N 48	82	73	75
..	...	...	...	...	...	...
195	196	Tirta Ramadhani	P SD N 23	54	69	82
196	197	Gita Nugraha	P SD N 23	76	80	60
197	198	Citra Anggara	P SD N 23	99	66	56
198	199	Lala Ramadhani	L SD N 23	71	78	59
199	200	Fajar Kartika	L SD N 23	56	92	99

[200 rows x 7 columns]

Gambar 4. 28 Data Sampel Untuk Pengujian Prediksi Cluster Pada Data Baru



## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

```

## untuk mengecek nilai NaN pada data training
print("NaN dalam x_train:", np.isnan(x_train).sum())
print("NaN dalam x_test:", np.isnan(x_test).sum())

NaN dalam x_train: 0
(P/L)
Sekolah
Nilai 1
Nilai 2
Nilai 3
Cluster
Nilai Rata-Rata
dtype: int64
NaN dalam x_test: 0
(P/L)
Sekolah
Nilai 1
Nilai 2
Nilai 3
Cluster
Nilai Rata-Rata
dtype: int64

imputer = SimpleImputer(strategy='mean')
x_train = imputer.fit_transform(x_train)
x_test = imputer.transform(x_test)

print("NaN dalam x_train setelah preprocessing:", np.isnan(x_train).sum())
print("NaN dalam x_test setelah preprocessing:", np.isnan(x_test).sum())

NaN dalam x_train setelah preprocessing: 0
NaN dalam x_test setelah preprocessing: 0

```

Gambar 4. 30 Pengecekan Nilai Nan Data Training dan Mengisi Nilai Hilang

### 4.3.2.1.17. Hitung Nilai Akurasi K1-10

Menghitung nilai akurasi K terbaik dan akan dilihat nilai akurasi yang paling tinggi. Penghitungan nilai dapat dilihat pada gambar berikut:

```

# Loop untuk menghitung akurasi untuk setiap nilai k
accuracies = []
k_values = range(1, 11)

for k in k_values:
    knn = KNeighborsClassifier(n_neighbors=k)
    knn.fit(x_train, y_train)
    y_pred_k = knn.predict(x_test)
    acc = accuracy_score(y_test, y_pred_k)
    accuracies.append(acc)
    print(f"Akurasi untuk k={k}: {acc}")

Akurasi untuk k=1: 0.9509803921568627
Akurasi untuk k=2: 0.9117647058823529
Akurasi untuk k=3: 0.9411764705882353
Akurasi untuk k=4: 0.9215686274509803
Akurasi untuk k=5: 0.9313725490196079
Akurasi untuk k=6: 0.9019607843137255
Akurasi untuk k=7: 0.9117647058823529
Akurasi untuk k=8: 0.9117647058823529
Akurasi untuk k=9: 0.9411764705882353
Akurasi untuk k=10: 0.9215686274509803

```

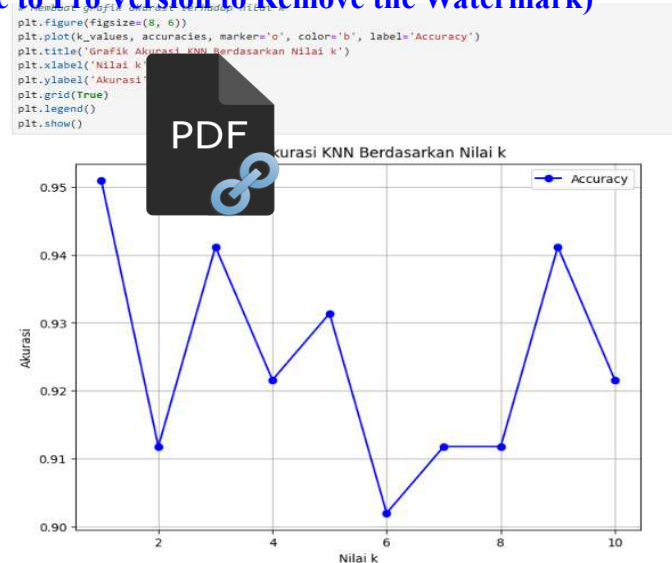
Gambar 4. 31 Nilai Akurasi K1-10

### 4.3.2.1.18. Grafik Accurasi *K-Nearest Neighbors (KNN)* Pada Nilai K

Menampilkan grafik visualisasi pada akurasi *K-Nearest Neighbors (KNN)* pada nilai K dari K1 sampai K10 . Akurasi knn dapat dilihat pada grafik gambar berikut:

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



Gambar 4. 32 Grafik Accurasi K-Nearest Neighbors (KNN) Pada Nilai K

### 4.3.2.1.19. Menampilkan Nilai K Terbaik

Dari grafik sudah terlihat Nilai akurasi k tertinggi terletak pada K1. Berikut dapat dilihat pada gambar ini:

```
# Mencetak nilai k dengan akurasi terbaik
best_k = k_values[np.argmax(accuracies)]
print(f"Nilai k terbaik adalah: {best_k} dengan akurasi {max(accuracies)}")
```

Nilai k terbaik adalah: 1 dengan akurasi 0.9509803921568627

Gambar 4. 33 Menampilkan Nilai K Terbaik

### 4.3.2.1.20. Pembuatan Model Confusion Matrik dan Accuracy Total

Pengujian Model Confusion Matrik dapat dilihat pada gambar berikut:

```
# prediksi data pengujian
AB.fit(x_train, y_train)
yab = AB.predict(x_train)

#Membuat model confusion matrix
clr= confusion_matrix(y_train,yab)
clr

array([[195,  0,  6],
       [ 4, 53,  0],
       [ 1,  0, 147]], dtype=int64)

#Menghitung akurasi total
acc_score = accuracy_score(y_train,yab)
acc_score

0.9729064039408867
```

Gambar 4. 34 Model Confusion Matrix dan Accuracy Total

## Protected by PDF Anti-Copy Free

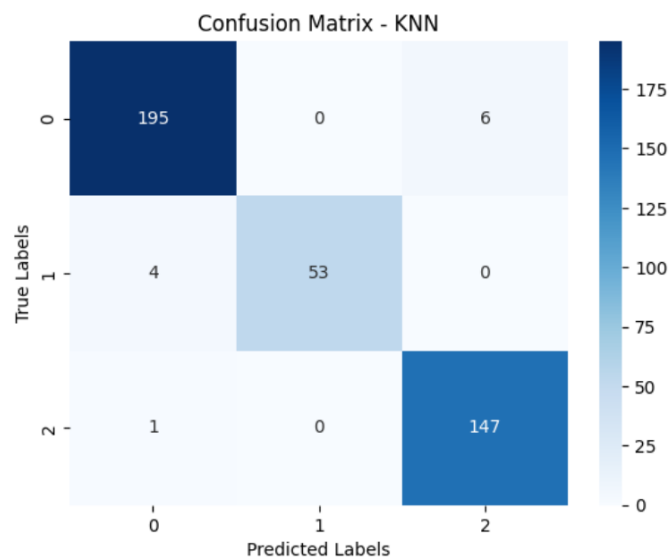
(Upgrade to Pro Version to Remove the Watermark)

### 4.3.2.1.21. Visualisasi Grafik Confusion Matrix

Mengukur keakuratan model klasifikasi, model melakukan klasifikasi dengan menunjukkan prediksi yang benar dan salah untuk setiap kelas. Grafik pengujian terlihat pada gambar berikut:

```
# menampilkan confusion matrix KNN
conf_matrix = confusion_matrix(y_train, yab)

sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues')
plt.title('Confusion Matrix - KNN')
plt.xlabel('Predicted Labels')
plt.ylabel('True Labels')
plt.show()
```



Gambar 4. 35 Visualisasi Grafik Confusion Matrix

### 4.3.2.1.22. Menghitung Akurasi, Precision, Recall Sensitivity, Specificity, FPR, FNR, dan F1-score.

Dengan menghitung metrik *precision*, *recall*, *specificity*, *FPR*, *FNR*, dan *F1-score*, mendapatkan gambaran yang lebih lengkap tentang kinerja model *KNN* dan membuat keputusan yang lebih baik dalam pengembangan dan penerapan model pembelajaran mesin.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

```
#Menghitung precision, recall, sensitivity, specificity, FPR, FNR, dan F1-score
cm=confusion_matrix(y_train.ravel(), yab.ravel())
tn = cm[0][0]
fn = cm[1][0]
tp = cm[1][1]
fp = cm[0][1]

pre = tp / (tp + fn)
rec = tp / (tp + fn)
spe = tn / (tn + fp)
fal = fp / (fp + tn)
mis = fn / (fn + tp)
f1 = (2*rec*pre) / (rec+pre)
print("Calculating precision for each class: ",pre)
print("Calculating recall (also called the true positive rate or the sensitivity) for each class: ",rec)
print("Calculating specificity (also called the true negative rate) for each class: ",spe)
print("Calculating fall out (also called the false positive rate) for each class :",fal)
print("Calculating miss rate (also called the false negative rate) for each class :",mis)
print("Calculating F1-Score for each class: ",f1)

Calculating precision for each class: 0.9298245614035088
Calculating recall (also called the true positive rate or the sensitivity) for each class: 0.9298245614035088
Calculating specificity (also called the true negative rate) for each class: 1.0
Calculating fall out (also called the false positive rate) for each class : 0.0
Calculating miss rate (also called the false negative rate) for each class : 0.07017543859649122
Calculating F1-Score for each class: 0.9298245614035088
```

Gambar 4. 36 Akurasi, Precision, Recal Sensitivity, Specificity, FPR, FNR, dan F1-score.

### 4.3.2.1.23. Menampilkan Accuracy, Precision, Recall dan F1-score

Berdasarkan output dari `classification_report` tersebut, kita dapat mengevaluasi kinerja model klasifikasi dengan *precision*, *recall*, *f1-score*, dan *accuracy*. Dapat kita lihat bahwa Model memiliki kinerja **sangat baik** dengan *akurasi keseluruhan 97%*.

*Model K-Nearest Neighbors (KNN)* bekerja secara optimal dalam mengenali pola dari data yang diberikan. Perbedaan kecil antara precision dan recall menunjukkan bahwa model tidak bias dan memiliki keseimbangan yang baik antara memprediksi positif dan negatif.

```
print(classification_report(y_train, yab))
```

	precision	recall	f1-score	support
0	0.97	0.97	0.97	201
1	1.00	0.93	0.96	57
2	0.96	0.99	0.98	148
accuracy			0.97	406
macro avg	0.98	0.96	0.97	406
weighted avg	0.97	0.97	0.97	406

Gambar 4. 37 Menampilkan Accuracy, Precision, Recall dan F1-score

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### 4.3.2.1.24. Save Data Baru Untuk Inputan Data Baru

Simpan data hasil klasifikasi digunakan untuk inputan data baru yang akan dibuat, data tersebut pada gambar berikut:

```
# Export the DataFrame to a CSV file
data.to_csv('clustered_data_knn1.csv', index=False) # Set index=False to avoid
```

```
data = pd.read_csv('clustered_data_knn1.csv')
data
```

	No.	Nama	(P/L)	Sekolah	Nilai 1	Nilai 2	Nilai 3	Cluster	D.Cluster
0	80	Nazella Aura	P	SD N 23	75	79	75	0	menengah
1	317	Qori Anisa	P	SD N 23	90	91	92	1	tinggi
2	486	Vina Rizki	P	SD N 23	89	86	85	1	tinggi
3	397	Iqbal Akbar	L	SD N 48	78	70	82	0	menengah
4	168	Sofia putri	P	SD N 48	79	77	80	0	menengah
...	...	...	...	...	...	...	...	...	...
503	107	Abdul Qodir	L	SD N 23	83	83	84	0	menengah
504	271	Dinda Safitri	P	SD N 23	93	91	94	1	tinggi
505	349	Kiki Lestari	P	SD N 48	90	92	91	1	tinggi
506	436	Vino Pranata	L	SD N 48	93	90	92	1	tinggi
507	103	Shoza izareen	P	SD N 23	70	65	71	2	rendah

Gambar 4. 38 Data Terbaru Hasil Klasifikasi

### 4.3.2.1.25. Input Data Baru

Input yang kedua ini adalah input data baru sebagai hasil akhir dari penelitian, berisikan input nama siswa, nilai calistung, rata rata nilai dan akan menghasilkan output siswa termasuk ke bagian nilai rendah, menengah atau tinggi. Hal ini dapat membantu proses perkelompokkan kemampuan calistung siswa agar dapat memberikan pembelajaran yang efektif sesuai dengan kemampuan siswa masing masing.

```
import pandas as pd
import numpy as np

# Fungsi untuk memprediksi cluster berdasarkan jarak ke centroid
def predict_cluster(average_score):
    # Centroid dari tiga cluster
    centroids = np.array([[0.61904762, 0.56470588, 0.52941176], # Cluster Menengah (0)
                          [0.87179487, 0.85234118, 0.85882353], # Cluster Tinggi (1)
                          [0.20238095, 0.21568627, 0.1952381]] # Cluster Rendah (2)

    # Menghitung jarak ke tiap centroid
    distances = np.linalg.norm(centroids - average_score, axis=1)
    predicted_cluster = np.argmin(distances)

    # Mapping hasil cluster ke kategori
    if predicted_cluster == 0:
        return 'Menengah'
    elif predicted_cluster == 1:
        return 'Tinggi'
    elif predicted_cluster == 2:
        return 'Rendah'
    else:
        return 'Unknown'

# Load data dari file CSV
data = pd.read_csv('clustered_data_knn1.csv')

# Hitung nilai rata-rata dari kolom Nilai 1, Nilai 2, dan Nilai 3
data['average_score'] = data[['Nilai 1', 'Nilai 2', 'Nilai 3']].mean(axis=1)

def predict_cluster_from_data(average_score):
    """
    Prediksi cluster berdasarkan dataset
    """
    # Cari baris dengan nilai rata-rata terdekat
    closest_row = data.iloc[(data['average_score'] - average_score).abs().idxmin()]

    # Ambil cluster dari baris tersebut
    cluster_number = closest_row['Cluster']
```

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

```

# Mapping angka cluster ke kategori
if cluster_number == 0:
    return 'Menengah'
elif cluster_number == 1:
    return 'Tinggi'
elif cluster_number == 2:
    return 'Rendah'
else:
    return 'Unknown'

# Input nama dari pengguna
nama = input("Masukkan nama Anda: ")

try:
    # Input nilai 1, 2, dan 3 dari pengguna
    nilai1 = float(input("Masukkan nilai 1: "))
    nilai2 = float(input("Masukkan nilai 2: "))
    nilai3 = float(input("Masukkan nilai 3: "))

    # Hitung rata-rata dari ketiga nilai
    average_score = (nilai1 + nilai2 + nilai3) / 3
    average_score_array = np.array([average_score, average_score, average_score])

    # Prediksi cluster berdasarkan centroid
    predicted_cluster = predict_cluster(average_score_array)

    # Tampilkan hasil
    print(f"\nNama: {nama}")
    print(f"Nilai rata-rata: {average_score:.2f}")

    # Prediksi cluster berdasarkan data CSV
    cluster_result = predict_cluster_from_data(average_score)
    print(f"Kategori cluster berdasarkan dataset: {cluster_result}")

except ValueError:
    print("Input harus berupa angka.")

```

Masukkan nama Anda:

Masukkan nama Anda: Amanda Salsabila

Masukkan nilai 1: 90

Masukkan nilai 2: 100

Masukkan nilai 3: 88

Nama: Amanda Salsabila

Nilai rata-rata: 92.67

Kategori cluster berdasarkan dataset: Tinggi

Gambar 4. 39 Hasil Input Akhir

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### BAB V KESIMPULAN DAN SARAN



#### 5.1 Kesimpulan

Berdasarkan pembahasan yang ada, maka dapat diambil kesimpulan penelitian yaitu dengan adanya sistem Metode hybrid dalam Pengelompokkan kemampuan calistung siswa berbasis machine learning maka, dua Sekolah Dasar Kota Lubuklinggau ini tidak lagi mengalami kesulitan untuk mengelompokkan kemampuan calistung siswa dikarenakan telah memiliki metode pengelompokkan untuk kemampuan calistung siswa berbasis machine learnig. Penelitian ini menunjukkan bahwa *clustering* menggunakan *K-Means* berhasil mengelompokkan siswa berdasarkan nilai mereka, yang dapat digunakan untuk analisis lebih lanjut dalam Pendidikan didapatkan 3 Cluster nilai yaitu Menengah(0), Rendah(1) dan Tinggi(2). Setelah itu, pada pengujian model klasifikasi *K-Nearest Neighbors (KNN)* dengan  $k=3$  yang memiliki akurasi tertinggi yaitu **95%** memberikan akurasi pada pengujian model klasifikasi *K-Nearest Neighbors (KNN)* yang sangat baik dengan **akurasi 97%**, **Precision cluster 0 sebesar 97%, cluster 1 sebesar 100%, cluster 2 sebesar 96%. Recall cluster 0 sebesar 97%, cluster 1 sebesar 93%, cluster 2 sebesar 99%. F1-Score cluster 0 sebesar 97%, cluster 1 sebesar 96%, cluster 2 sebesar 98%** dalam klasifikasi data siswa yng dikategorikan “sangat baik” .


#### 5.1 Saran

Hasil penelitian ini dapat digunakan sebagai dasar untuk pengembangan lebih lanjut dalam analisis data pendidikan dan penerapan teknik machine learning dalam konteks yang lebih luas, dalam dataset yang akan diolah harus lebih banyak lagi agar akurasi yang dihasilkan akan lebih efektif dan pembuatan sistem aplikasi agar lebih mempermudah penggunaan sistem pngelompokan kemampuan calistung siswa untuk sekolah dan guru guru.

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

### DAFTAR PUSTAKA

- 
- [1] A. Candra Dewi *et al*, “Pengaruh Kemajuan Teknologi dalam Dunia Pendidikan,” *J. Educ.*, vol. 1, no. 01, pp. 9725–9734, 2023.
- [2] P. Weraman, I. P. Agus, and D. Hita, “KEMAMPUAN CALISTUNG DAN MOTIVASI BELAJAR,” vol. 4, no. 2, pp. 1234–1239, 2023.
- [3] L. Latifah and F. P. Rahmawati, “Penerapan Program CALISTUNG untuk Meningkatkan Literasi Numerasi Siswa Kelas Rendah di Sekolah Dasar,” *J. Basicedu*, vol. 6, no. 3, pp. 5021–5029, 2022, doi: 10.31004/basicedu.v6i3.3003.
- [4] J. Jumrodah, M. Oktaviany, D. Amelia Safitri, U. Amalia, Y. Yunita, and L. Safitri, “Penerapan Model Pembelajaran Calistung di Rumah Pintar Sebagai Upaya Pendidikan Anak-Anak di Desa Walur,” *J. Pengabd. Kpd. Masy. Nusant.*, vol. 5, no. 2, pp. 2442–2450, 2024, doi: 10.55338/jpkmn.v5i2.3201.
- [5] A. Safitri, H. Y. Ayyasy, H. Purba, M. B. Winanda, and ..., “Upaya Peningkatan Kemampuan Calistung Siswa Kelas I Sd Negeri 106224 Desa Kerapuh,” *Pedamas ...*, vol. 2, pp. 293–299, 2024, [Online]. Available: <https://pekatpkm.my.id/index.php/JP/article/view/219%0Ahttps://pekatpkm.my.id/index.php/JP/article/download/219/193>
- [6] S. Wahyuni, S. Syamsuyurnita, D. Kesuma, and ..., “Peningkatan Kemampuan Literasi Numerasi Melalui Program Calistung Di Sanggar Bimbingan Kampung Bharu Kuala Lumpur, Malaysia,” *Innov. J. ...*, vol. 3, pp. 1652–1662, 2023, [Online]. Available: <http://j-innovative.org/index.php/Innovative/article/view/3669%0Ahttps://j-innovative.org/index.php/Innovative/article/download/3669/2633>
- [7] L. Juliasih, G. Sya’ban Maulana, S. Rahmawati, and L. Siti Muslimah, “Pendampingan Siswa Melalui Metode Calistung di Desa Mekarsari Kecamatan Mekarmukti Kabupaten Garut,” *Bantenese J. Pengabd. Masy.*, vol. 6, no. 1, pp. 1–8, 2024, doi: 10.30656/ps2pm.v6i1.7782.
- [8] R. A. Ananda, “Clustering Menggunakan Algoritma K-Means untuk Mengelompokan Data Perjudian Berdasarkan Wilayah di Kota Binjai ( Studi

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

Kasus : Pengadilan Negeri Binjai ),” no. 4, 2024.

- [9] H. Hairani and H. A. Id..., “Pengaruh Implementasi Machine Learning Pada Bidang Pendidikan,” *J. PDF dan Pemberdaya. Masy. Tahun*, vol. 2, no. 2, pp. 305–310, 2022, doi: 10.30821/adma.v2i2.xxxx.
- [10] A. Asmana, Y. Arie Wijaya, and M. Martanto, “Clustering Data Calon Siswa Baru Menggunakan Metode K-Means Di Sekolah Menengah Kejuruan Wahidin Kota Cirebon,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 552–559, 2022, doi: 10.36040/jati.v6i2.5236.
- [11] Dewi Eka Putri and Eka Praja Wiyata Mandala, “Hybrid Data Mining berdasarkan Klasterisasi Produk untuk Klasifikasi Penjualan,” *J. KomtekInfo*, vol. 9, pp. 68–73, 2022, doi: 10.35134/komtekinfo.v9i2.279.
- [12] A. E. Wibowo and T. Habanabakize, “K-Means Clustering untuk Klasifikasi Standar Kualifikasi Pendidikan dan Pengalaman Kerja Guru SMK di Indonesia,” *J. Din. Vokasional Tek. Mesin*, vol. 7, no. 2, pp. 152–163, 2022, doi: 10.21831/dinamika.v7i2.53848.
- [13] T. S. S. W. I. K. V. Ramadhan, “Teori Behaviourisme Menurut Pavlov Dalam Pembelajaran Permulaan Calistung,” *J. Ilm. Pendidik. Dasar*, vol. 09, 2024.
- [14] H. Wulandari and H. A. Azizah, “Penerapan Calistung di PAUD,” *JECIE (Journal Early Child. Incl. Educ.*, vol. 7, no. 1, pp. 11–21, 2023, doi: 10.31537/jecie.v7i1.1426.
- [15] J. Multidisiplin *et al.*, “Peran mahasiswa dalam meningkatkan kemampuan calistung pada siswa kelas 2 sdn 1 desa bumi kencana dengan media flashcard,” vol. 1, no. 3, pp. 429–434, 2024.
- [16] D. Anzelina, K. S. Partawan, and ..., “Pendampingan Program Calistung Pada Siswa Sekolah Dasar di Desa Sigaol Marbun,” *Int. J. ...*, vol. 2, no. 2, pp. 1–5, 2023, [Online]. Available: <https://amwiinstitute.id/index.php/ijecs/article/view/21%0Ahttps://amwiinsitute.id/index.php/ijecs/article/download/21/14>
- [17] S. Apriyanti and E. Aprianti, “Dampak penyelenggaraan aktivitas baca, tulis dan hitung (Calistung) pada anak usia dini,” *J. CERIA (Cerdas Energik*

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

- Responsif Inov. Adapt.*, vol. 6, no. 4, pp. 399–407, 2023, [Online]. Available: <https://journal.ikipsiliwangi.ac.id/index.php/ceria/article/view/17484>
- [18] M. Mulyatno, A. Pujitomo, D. K. Legowo, A. Firman, and A. R. Mahendra, “Pemberdayaan Guru Sekolah Menengah Kejuruan Melalui Pelatihan Pengenalan Machine Learning,” *J. Pengabd. Masy. Bangsa*, vol. 1, no. 11, pp. 2899–2904, 2024, doi: 10.59837/jpmba.v1i11.628.
- [19] N. Hilmizen, A. Munandar, J. Muryati, and A. Mulyanto, “Analisa Peningkatan Softskill dan Hardskill Siswa melalui Kurikulum Merdeka dengan Machine Learning,” *EduInovasi J. Basic Educ. Stud.*, vol. 4, no. 2, pp. 306–319, 2024, doi: 10.47467/edu.v4i2.1582.
- [20] A. Fathurohman, “Machine Learning Untuk Pendidikan: Mengapa Dan Bagaimana,” *J. Inform. dan Teknol. Komput.*, vol. 1, no. 3, pp. 57–62, 2021, [Online]. Available: <https://journal.amikveteran.ac.id/index.php/jitek/article/view/306>
- [21] N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, “Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning,” *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 1, pp. 47–58, 2021, doi: 10.28932/jutisi.v7i1.3216.
- [22] M. Pandia, “Kajian Literatur Multimedia Retrieval : Machine Learning Untuk Pengenalan Wajah,” *J. Ilmu Komput. dan Sist. Inf.*, vol. 7, no. 1, pp. 161–166, 2024, doi: 10.55338/jikomsi.v7i1.2758.
- [23] F. A. Astuti, “Pemanfaatan Teknologi Artificial Intelligence untuk Penguatan Kesehatan dan Pemulihan Ekonomi Nasional,” *J. Sist. Cerdas*, vol. 4, no. 1, pp. 25–34, 2021, doi: 10.37396/jsc.v4i1.124.
- [24] S. Asyuti and A. A. Setyawan, “Data Mining Dalam Penggunaan Presensi Karyawan Denga Cluster Means,” *J. Ilm. Sains Teknol. Dan Inf.*, vol. 1, no. 1, pp. 01–10, 2023, [Online]. Available: <https://jurnal.alimspublishing.co.id/index.php/JITI/article/download/6/6>
- [25] C. Zai, “Implementasi Data Mining Sebagai Pengolahan Data,” *J. Portal Data*, vol. 2, no. 3, pp. 1–12, 2022, [Online]. Available: <http://portaldata.org/index.php/portaldata/article/view/107>

## Protected by PDF Anti-Copy Free

[\(Upgrade to Pro Version to Remove the Watermark\)](#)

- [26] Rini Rahmawati and Agus Bahtiar, “Pengelompokan Remaja Berdasarkan Segmentasi Usia Menggunakan Metode K-Means Clustering (Studi Kasus : Desa Sindangsari),” *AKUNTANSI*, vol. 2, no. 2, pp. 35–51, 2023, doi: 10.55606/akuntansi.v2i2
- [27] I. F. Tarigan, D. Hartama, and I. S. Saragih, “Penerapan Data Mining Pada Prediksi Kelayakan Pemohon Kredit Mobil Dengan K-Medoids Clustering,” *KLIK Kaji. Ilm. Inform. ...*, vol. 1, no. 4, pp. 170–179, 2021, [Online]. Available: <http://www.djournals.com/klik/article/view/153>
- [28] D. K. Damanik, A. Antoni, and M. Z. Siambaton, “Algoritma K-Means untuk Klasterisasi Obat di Puskesmas Simpang Limun Kota Medan,” *Hello World J. Ilmu Komput.*, vol. 3, no. 2, pp. 100–110, 2024, doi: 10.56211/helloworld.v3i2.549.
- [29] D. Herwanis and Susidamaiyanti, “Peningkatan Kemampuan Menulis Teks Deskriptif Berbahasa Inggris Melalui Teknik Clustering Pada Program Studi Tadris Bahasa Inggris Di Stain Gajah Putih Takengon,” *J. As-Salam*, vol. 4, no. 2, pp. 216–230, 2020, doi: 10.37249/as-salam.v4i2.185.
- [30] Y. Filki, “Algoritma K-Means Clustering dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Dana Desa,” *J. Inform. Ekon. Bisnis*, vol. 4, pp. 166–171, 2022, doi: 10.37034/infeb.v4i4.166.
- [31] R. Fauziah and A. I. Purnamasari, “Implementasi Algoritma K-Means pada Kasus Kekerasan Anak dan Perempuan Berdasarkan Usia,” *Hello World J. Ilmu Komput.*, vol. 2, no. 1, pp. 34–41, 2023, doi: 10.56211/helloworld.v2i1.232.
- [32] F. Andini, D. Zilfitri, Y. Filki, and M. Ridho, “Algoritma K-Means Clustering dalam Optimalisasi Komposisi Pakan Ternak Ayam Petelur,” *J. Sistik Inf. dan Teknol.*, vol. 5, pp. 44–48, 2022, doi: 10.37034/jsisfotek.v5i2.168.
- [33] U. Muslim and N. Al Washliyah, “Prosiding Seminar Nasional Hasil Pengabdian 2022 HYBRID LEARNING SEBAGAI METODE PEMBELAJARAN LITERASI DI MASA PANDEMI Fita Fatria 1), Tiflatul Husna 2),” no. 1, pp. 83–88, 2022, [Online]. Available:

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

<https://sevima.com/apa-perbedaan-blended-learning-dan->

- [34] A. Premana, O. Saeful and A. Pandhu Wijaya, “Klasifikasi Jenis Mangga Apel Menggunakan Metode K-Means Klustering,” *J. Tek. Indones.*, vol. 1, no. 1, pp. 16–25, 2024, doi:10.58860/jti.v1i1.1.
- [35] Y. Bagus Pratama and A. Setiawan, “Implementasi Machine Learning Menggunakan Algoritma K-Means Untuk Klasifikasi Sekolah Dasar,” *RESOLUSI Rekayasa Tek. Inform. dan Inf.*, vol. 4, no. 3, pp. 249–257, 2024, [Online]. Available: <https://djournals.com/resolusi>

**Protected by PDF Anti-Copy Free**  
 (Upgrade to Pro Version to Remove the Watermark)  
 LAMPIRAN

Berikut ini adalah Lampiran dan referensi yang mendukung data dan informasi yang disajikan dalam proposal ini :

**YAYASAN PENDIDIKAN DWI TUNGGAL PALEMBANG**  
**UNIVERSITAS BINA INSAN**

**Formulir Pengajuan Judul Skripsi**  
**Program Studi Informatika**

Nama : Amanda Salsabila  
 NIM : 2102020001  
 Alamat : Jl. Nangka Kacung, Rt.03 No.77 Kel.Ponorogo, Kec.Lubuklinggau Utara II  
 No.Hp : 083171117010

Rumusan Masalah 1 : Bagaimana metode hybrid machine learning dapat diterapkan untuk mengetahui kelompok kemampuan Calistung siswa?  
 Judul 1 : Metode Hybrid dalam Pengelompokan Kemampuan Calistung Siswa Berbasis Machine Learning.

Rumusan Masalah 2 : Bagaimana menerapkan metode data mining menggunakan decision trees untuk mengklasifikasikan jenis layanan berdasarkan data pemesanan di Studio Semangko Photography?  
 Judul 2 : Implementasi Data Mining untuk Mengkalsifikasi Layanan dari Data Booking di Studio Foto Menggunakan Metode decision trees (Studi case : Studio Semangko Photgrafy)

Rumusan Masalah 3 : Bagaimana cara menerapkan algoritma Naive Bayes untuk mendeteksi sentimen pengguna terhadap aplikasi e-learning (Studi kasus: Brainly dan Quizizz) dan mengetahui aplikasi mana yang memiliki sentimen pengguna lebih baik?  
 Judul 3 : Implementasi Algoritma Naive Bayes untuk Mendeteksi Sentimen Pengguna Aplikasi E-Learning (Studi Kasus: Brainly dan Quizizz).

**Diusulkan Judul Nomor :1(satu)/ 2(Dua)/ 3(Tiga)\***

Lubuklinggau, 25 September 2024  
 Mahasiswa yang mengusulkan,

(Amanda Salsabila)

Menyetujui Dosen Pembimbing,  
 Pembimbing 1 (Andri Anto Tri Susilo, M.kom) .....  
 Pembimbing 2 (Nelly Khaiani Daulay, M.Kom) .....

**Mengesahkan,**  
 Dekan Fakultas Ilmu Teknik  
 (Dr. Rudi Kurniawan, S.T., M.Kom.,)


**Mengetahui,**  
 Ketua Program Studi,  
 (Budi Santoso S.Kom., M.Kom.)

0733-4553932 (Rektorat Universitas Bina Insan) 0812-1826-6228 (Marketing UNIVBI) 0733-3280300 0852-3151-5800 (Admin UNIVBI)

Lampiran 1. Judul Penelitian

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



**UNIVERSITAS BINA INSAN**  
 FAKULTAS ILMU TEKNIK  
 PROGRAM STUDI INFORMATIKA

Jalan Jendral Binsar  
 Lubuklinggau, Kota Lubuklinggau, Provinsi Sumatera Selatan

Nomor : 0947c/UN 2024  
 Lampiran : -  
 Perihal : Permohonan

Kepada Yth.  
**Kepala Sekolah SD Negeri 23 Kota Lubuklinggau**  
 di-  
 Tempat

***Assalamualaikum Warahmatullahi Wabarakatuh***


Kami dari Fakultas Ilmu Teknik Universitas Bina Insan Lubuklinggau, dengan ini mengajukan permohonan kepada Bapak/Ibu untuk dapat kiranya menerima Mahasiswa kami berikut ini:

Nama : Amanda Salsabila  
 NIM : 2102020001  
 Jenjang Pendidikan : Strata 1 (S1)  
 Fakultas : Ilmu Teknik  
 Program Studi : Informatika  
 Judul Penelitian : Metode Hybrid Dalam Pengelompokan Kemampuan Calistung Siswa Berbasis Machine Learning

Penelitian ini hanya di lakukan untuk kepentingan akademik dan penulisan Karya tulis ilmiah (Skripsi) mahasiswa. Atas bantuan Bapak/Ibu kami ucapkan terima kasih.

***Wassalamualaikum Warahmatullahi Wabarakatuh***

Lubuklinggau, 1 Oktober 2024  
 Dekan Fakultas Ilmu Teknik,

  
**Dr. Rudi Kurniawan, ST.,M.Kom**

Tembusan Yth.  
 1. Ketua Yayasan Pendidikan Dwi Tunggal Palembang (sebagai laporan)  
 2. Rektor Universitas Bina Insan Lubuklinggau (sebagai laporan)  
 3. Arsip

0733-4553932 (Rektorat Universitas)    0812-1826-6228 (Marketing UNIVEBI)  
 0733-3280300 (Bina Insan)    0852-3151-5800 (Admin UNIVEBI)  
 0733-3380200 (Pusat Studi IT)

Lampiran 2. Surat Permohonan Izin Penelitian

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)

**UNIVERSITAS BINA INSAN**  
FAKULTAS ILMU TEKNIK  
PROGRAM STUDI INFORMATIKA

Nomor : 0947c/UN/2024  
Lampiran : -  
Perihal : Permohonan Izin Penelitian

Kepada Yth.  
**Kepala Sekolah SD Negeri 48 Kota Lubuklinggau**  
di-  
Tempat

*Assalamualaikum Warahmatullahi Wabarakatuh*


Kami dari Fakultas Ilmu Teknik Universitas Bina Insan Lubuklinggau, dengan ini mengajukan permohonan kepada Bapak/Ibu untuk dapat kiranya menerima Mahasiswa kami berikut ini:

Nama : Amanda Salsabila  
NIM : 2102020001  
Jenjang Pendidikan : Strata 1 (S1)  
Fakultas : Ilmu Teknik  
Program Studi : Informatika  
Judul Penelitian : Metode Hybrid Dalam Pengelompokan Kemampuan Calistung Siswa Berbasis Machine Learning

Penelitian ini hanya di lakukan untuk kepentingan akademik dan penulisan Karya tulis ilmiah (Skripsi) mahasiswa. Atas bantuan Bapak/Ibu kami ucapkan terima kasih.

*Wassalamualaikum Warahmatullahi Wabarakatuh*

Lubuklinggau, 1 Oktober 2024  
Dekan Fakultas Ilmu Teknik,

  
UNIVERSITAS BINA INSAN  
FAKULTAS ILMU TEKNIK

**Dr. Rudi Kurniawan, ST.,M.Kom**

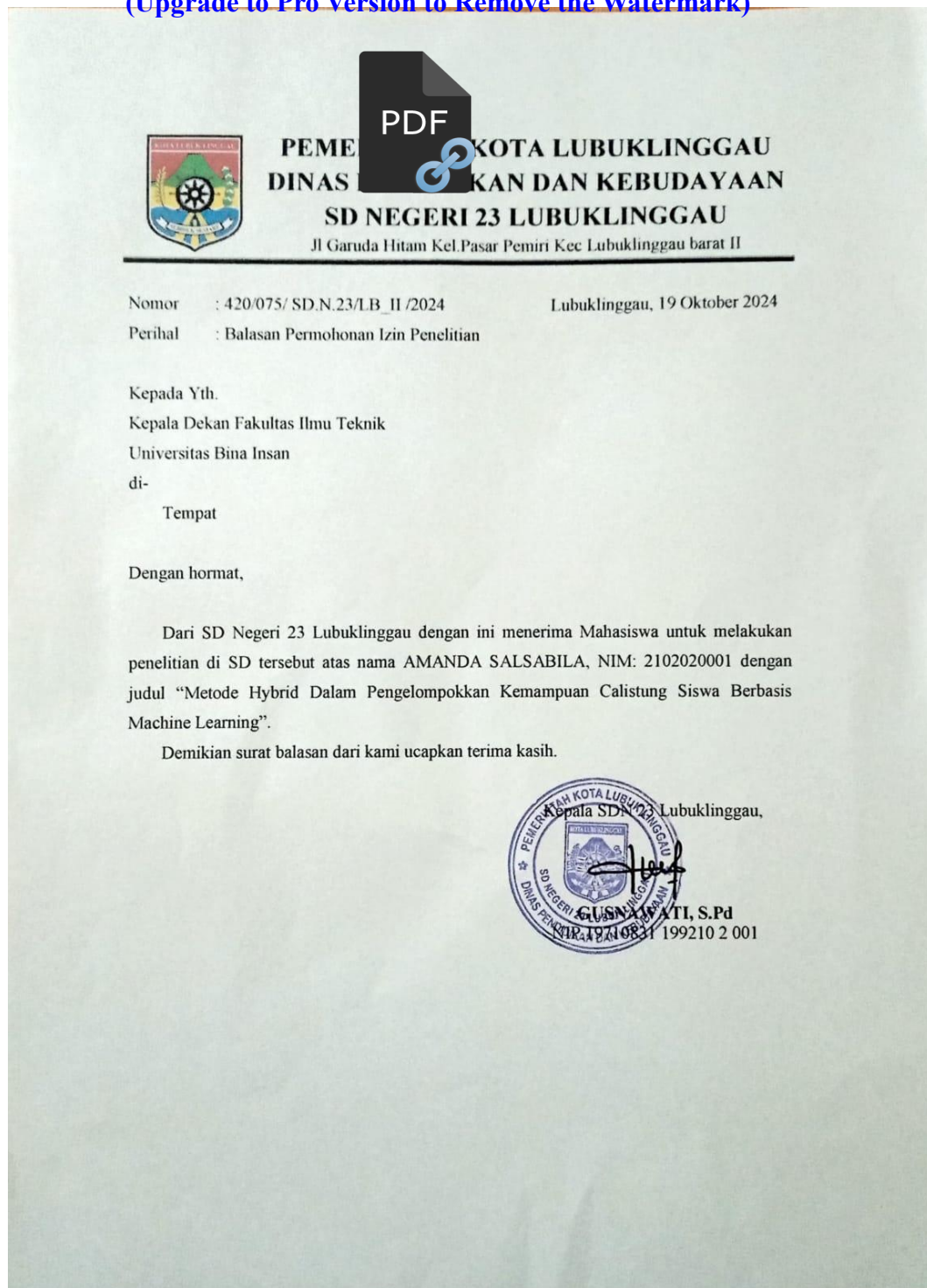
Tembusan Yth.  
1. Ketua Yayasan Pendidikan Dwi Tunggal Palembang (sebagai laporan)  
2. Rektor Universitas Bina Insan Lubuklinggau (sebagai laporan)  
3. Arsip

0733-4553932 (Rektorat Universitas Bina Insan)  
0733-3280300 (Bina Insan)  
0733-3280200 (Desa Cakrawala)

0812-1826-6228 (Marketing UNIVBI)  
0852-3151-5800 (Admin UNIVBI)

Lampiran 3. Surat Permohonan Izin Penelitian

**Protected by PDF Anti-Copy Free**  
 (Upgrade to Pro Version to Remove the Watermark)




Lampiran 4. Surat Balasan Permohonan Izin Pelitian



## Protected by PDF Anti-Copy Free




(Upgrade to Pro Version to Remove the Watermark)



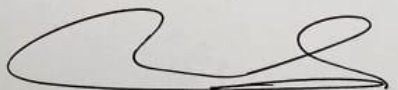
**UNIVERSITAS BINA INSAN**  
FAKULTAS ILMU TEKNIK

**LEMBAR PEMBIMBINGAN PROPOSAL SKRIPSI**

Nama : Amanda Sa  
 Nim : 210202001  
 Program Studi : Informatika  
 Pembimbing 1 : Andri Anto Tri Susilo, M.kom.  
 Pembimbing 2 : Nelly Khaiani Daulay, M.Kom  
 Judul : Metode Hybrid dalam Pengelompokkan Kemampuan Calistung Siswa Berbasis Machine Learning

NO	TANGGAL	TOPIK	KOMENTAR PEMBIMBING	TANDA TANGAN PEMBIMBING	
				1	2
1.	13 / 2024 / 11		<ul style="list-style-type: none"> <li>- Perulisan pada proposal harus di-</li> <li>- Perhatikan lagi dan perbaiki</li> <li>- Perjelasan algoritma yang akan di-</li> <li>- pakai 1 atau 2.</li> </ul>		
2.	15 / 2024 / 11		<ul style="list-style-type: none"> <li>- Perhatikan pada kata asing harus</li> <li>- gunakan italic atau miring.</li> <li>- Perbaiki sub bab agar lebih rapi</li> <li>- Rata Kiri kanan pada tabel hasil</li> <li>- penelitian terdahulu.</li> <li>- Tabel dibuat 1,2,3 jangan gunakan</li> <li>- 1.1</li> </ul>		
3.	19 / 2024 / 11		Acc, silakan lanjut proposal.		

Lubuklinggau, 2024  
Ketua Program Studi Ilmu Teknik


  
(Budi Santoso S.Kom., M.Kom)

0733-4553932 (Rektorat Universitas)    0812-1026-6228 (Marketing UNIVBI)  
 0733-3280300 (Bina Insan)    0652-3151-5800 (Admin UNIVBI)  
 0733-3280200 (Pascasarjana)    Admin@univbinainsan.ac.id    univbinainsan.ac.id - pascq.univbinainsan.ac.id

Lampiran 6. Lembar Bimbingan Proposal Dosen Pembimbing I

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



YAYASAN PENDIDIKAN DWI TUNGGAL PALEMBANG  
UNIVERSITAS BINA INSAN  
FAKULTAS ILMU TEKNIK  
Jalan Jendral Sudirman No. 13 Kel. Lubuk Plupung Kec. Lubuklinggau Selatan 1 Kota Lubuklinggau Provinsi Sumatera Selatan

**LEMBAR BIMBINGAN PROPOSAL SKRIPSI**

Nama : Amanda S  
Nim : 2102020001  
Program Studi : Informatika  
Pembimbing 1 : Andri Anto Tri Susilo, M.kom.  
Pembimbing 2 : Nelly Khaiani Daulay, M.Kom  
Judul : Metode Hybrid dalam Pengelompokkan Kemampuan Calistung Siswa Berbasis Machine Learning

NO	TANGGAL	TOPIK	KOMENTAR PEMBIMBING	TANDA TANGAN PEMBIMBING	
				1	2
1		Bab I	- Permasalahan sudah jelas - Gambaran metode yg digunakan ada - Data sampel harus jelas ditulis berapa jumlahnya		<i>[Signature]</i>
2.		BAB II	- Perhatikan typo pada tulisan - kerangka berpikir sudah baik dan jelas		<i>[Signature]</i>
3.		BAB III	- uraian Alat dan Bahan sebaiknya dibarengi narasi sesuai fungsi tabel		<i>[Signature]</i>
4.		BAB IV			
5.		Referensi	Gunakan referensi 5 Pahlawan terakir		<i>[Signature]</i>

Lubuklinggau, 2024  
Ketua Program Studi Ilmu Teknik


*[Signature]*  
(Budi Santoso S.Kom., M.Kom)

0733-4553932 (Rektorat Universitas)
0812-1826-6228 (Marketing UNIVBI)
0733-3280300 Bina Insan
0852-3151-5800 (Admin UNIVBI)
0733-3280200 (Pascasarjana)
Admin@univbinainsan.ac.id
univbinainsan.ac.id - pasca.univbinainsan.ac.id

Lampiran 7. Lembar Bimbingan Proposal Dosen Pembimbing II

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



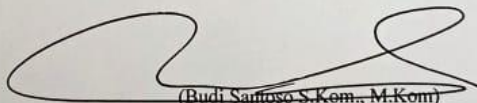
YAYASAN PENDIDIKAN DWI TUNGGAL PALEMBANG  
**UNIVERSITAS BINA INSAN**  
 FAKULTAS ILMU TEKNIK  
Jalan Jendral Sudirman No. 11 Kel. Lubuk Kuning Kec. Lubuklinggau Selatan 1 Kota Lubuklinggau Provinsi Sumatera Selatan

**LEMBAR BIMBINGAN PROPOSAL SKRIPSI**

Nama : Amanda S  
 Nim : 210202000  
 Program Studi : Informatika  
 Pembimbing 1 : Andri Anto Tri Susilo, M.kom.  
 Pembimbing 2 : Nelly Khaiani Daulay, M.Kom  
 Judul : Metode Hybrid dalam Pengelompokkan Kemampuan Calistung Siswa Berbasis Machine Learning

NO	TANGGAL	TOPIK	KOMENTAR PEMBIMBING	TANDA TANGAN PEMBIMBING	
				1	2
6.		Ace	Ace siap dilan zut leu Pembimbinga lee pembimbing ↑		d

Lubuklinggau, 2024  
 Ketua Program Studi Ilmu Teknik


  
 (Budi Santoso S.Kom., M.Kom)

☎ 0733-4553932 (Rektorat Universitas)
📞 0812-1826-6228 (Marketing UNIVBI)
📞 0733-3280300 (Bina Insan)
📞 0852-3151-5800 (Admin UNIVBI)
📧 0733-3280200 (Pascasarjana)
✉ Admin@univbinaisan.ac.id
🌐 univbinaisan.ac.id + pasca.univbinaisan.ac.id

Lampiran 8. Lembar Bimbingan Proposal Dosen Pembimbing II

## Protected by PDF Anti-Copy Free



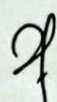
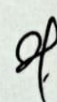
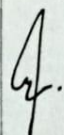
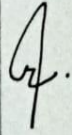
(Upgrade to Pro Version to Remove the Watermark)



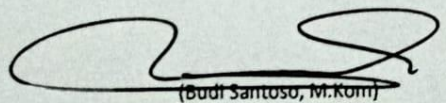
YAYASAN PENDIDIKAN DWI TUNGGAL PALEMBANG  
**UNIVERSITAS BINA INSAN**  
 FAKULTAS ILMU TEKNIK  
Jalan Jendral Besuki No. 100, Lubuklinggau Selatan 1 Kota Lubuklinggau Provinsi Sumatera Selatan

### LEMBAR PERBAIKAN SEMINAR PROPOSAL SKRIPSI

Nama Mahasiswa : Amanda Salsabila  
 NIM : 2101010001  
 Jenjang Pendidikan : Strata 1 ( S1 )  
 Fakultas : Ilmu Teknik  
 Program Studi : Informatika  
 Konsentrasi : -  
 Judul : Metode Hybird Dalam Pengelompokkan Kemampuan Calistung Siswa Berbasis Machine Learning

No	Dosen Penguji	Komentar Perbaikan	Tanda Tangan Ujian	Tanda Tangan Revisi
1	Andri Anto Tri Susilo, M.kom			
2	Melly Khairani Dhaulay, M.kom			
3	Armanto, M.kom	- Hitungan Manual		

Lubuklinggau, 14 Desember 2024  
 Ketua Program Studi Informatika


  
 (Budi Santoso, M.kom)

0733-4553932 (Rektorat Universitas)    0812-1826-6228 (Marketing UNIVBI)  
 0733-3280300 (Bina Insan)    0852-3151-5800 (Admin UNIVBI)  
 0733-3280200 (PascaSarjana)    Admin@univbinainsan.ac.id    univbinainsan.ac.id - pasca.univbinainsan.ac.id


Lampiran 9. Lembar Perbaikan Seminar Proposal Skripsi

## Protected by PDF Anti-Copy Free

(Upgrade to Pro Version to Remove the Watermark)



YAYASAN PENDIDIKAN DWI TUNGGAL PALEMBANG  
**UNIVERSITAS BINA INSAN**  
 FAKULTAS ILMU TEKNIK  
Jalan Jendral Sudirman No. 13 Kel. Lubuk Kumpang Kec. Lubuklinggau Selatan I Kota Lubuklinggau Provinsi Sumatera Selatan

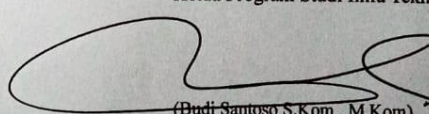


**BIMBINGAN SKRIPSI**

Nama : Amanda Salsabila  
 Nim : 2102020001  
 Program Studi : Informatika  
 Pembimbing 1 : Andri Anto Tri Susilo, M.kom.  
 Pembimbing 2 : Nelly Khairani Daulay, M.Kom  
 Judul : Metode Hybrid Dalam Pengelompokkan Kemampuan Calistung Siswa Berbasis Machine Learning

NO	TANGGAL	TOPIK	KOMENTAR PEMBIMBING	TANDA TANGAN PEMBIMBING	
				1	2
	27 / 2024 Desember	Bob 1-5	<ul style="list-style-type: none"> <li>- fitur rata rata pada dataset dihapus atau dihilangkan.</li> <li>- Perhatikan lagi alur pada pemrograman mem dibuat sesuai alur yang benar</li> <li>- Normalisasi pada preprocessing tidak harus dipakai.</li> </ul>	Am.	
	30 / 2024 12		<ul style="list-style-type: none"> <li>- Perbaiki proses pada klasifikasi KNN</li> </ul>	Am.	
	6 / 2025 1		<ul style="list-style-type: none"> <li>- Hasil prediksi data baru dijadikan satu dalam bentuk data frame.</li> </ul>	Am.	
	8 / 2025 1		Acc, Silakan uton <del>data</del> skripsi	Am.	

Lubuklinggau, 08 Januari 2024  
 Ketua Program Studi Ilmu Teknik


  
 (Budi Santoso S.Kom., M.Kom)

0733-4553932 (Rektorat Universitas)    0812-1826-6228 (Marketing UNIVBI)  
 0733-3280300 (Bina Insan)    0852-3151-5800 (Admin UNIVBI)  
 0733-3280200 (Pascasarjana)    ✉ Admin@univbinainan.ac.id    univbinainan.ac.id | pda.univbinainan.ac.id


Lampiran 10. Lembar Bimbingan Skripsi Pembimbing I

## Protected by PDF Anti-Copy Free

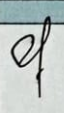
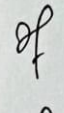
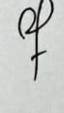
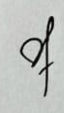
(Upgrade to Pro Version to Remove the Watermark)



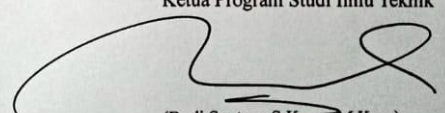
YAYASAN PENDIDIKAN DWI TUNGGAL PALEMBANG  
**UNIVERSITAS BINA INSAN**  
 FAKULTAS ILMU TEKNIK  
Jalan Sempura 113 Kel. Lubuk Kumpang Kec. Lubuklinggau Selatan I Kota Lubuklinggau Provinsi Sumatera Selatan



Nama : Amanda Salsabila  
 Nim : 2102020001  
 Program Studi : Informatika  
 Pembimbing 1 : Andri Anto Tri Susilo, M.kom.  
 Pembimbing 2 : Nelly Khairani Daulay, M.Kom  
 Judul : Metode Hybrid Dalam Pengelompokkan Kemampuan Calistung Siswa Berbasis Machine Learning

NO	TANGGAL	TOPIK	KOMENTAR PEMBIMBING	TANDA TANGAN PEMBIMBING	
				1	2
1.	20/2024 12		Perbaiki sambar ketangkasan berpikir.		
2.			Setiap rumus berikan penomoran		
3.			Kesimpulan buat dalam satu paragraf saja		
4.			Acc lanjut pembimbing 1		

Lubuklinggau, 08 Januari 2024  
 Ketua Program Studi Ilmu Teknik



(Bedi Santoso S.Kom., M.Kom)

0733-4553932 (Rektorat Universitas)    0612-1826-6228 (Marketing UNIVBI)  
 0733-3280300 (Bina Insan)    0652-3151-5800 (Admin UNIVBI)  
 0733-3280200 (Pascasarjana)    Admin@univbinainsan.ac.id    univbinainsan.ac.id - pasca.univbinainsan.ac.id

Lampiran 11. Lembar Bimbingan Skripsi Pembimbing II

**Protected by PDF Anti-Copy Free**  
(Upgrade to Pro Version to Remove the Watermark)



Lampiran 12. Dokumentasi Bersama Kepala Sekolah SD



Lampiran 13. Dokumentasi Pengambilan Data Nilai Siswa