

Mengukur Tingkat Literasi Anak Menggunakan Algoritma *Decision Tree* Dan *Logistic Regression*

Faris Gymnatsiar Rahman¹, Rizal Rachman²

¹Program Studi Teknik Informatika, Universitas Adhirajasa Reswara Sanjaya

²Program Studi Sistem Informasi, Universitas Adhirajasa Reswara Sanjaya

e-mail: ¹farisgymnatsiarrahmann@gmail.com, ²rizalrachman@ars.ac.id

Abstrak

Penelitian ini bertujuan untuk menganalisis kinerja siswa SDN Cilampeni pada mata pelajaran Teknologi Informasi dan Komunikasi (TIK). Data dikumpulkan dari 302 siswa, dengan 10 variabel penilaian yang mencakup wawasan pengetahuan, matematika komputer, Bahasa Inggris dengan komputer, menggambar dengan komputer, pengolahan kata, tugas, minat belajar, kehadiran, tingkah laku, serta tes praktek dan teori. Analisis data menunjukkan bahwa semua siswa berhasil lulus dengan nilai rata-rata yang tinggi di semua kategori. Hasil analisis menggunakan algoritma *Decision Tree* menunjukkan bahwa variabel kehadiran dan minat belajar adalah faktor penentu utama yang mempengaruhi kelulusan siswa, dengan akurasi model sebesar 95%. Sementara itu, analisis menggunakan algoritma *Logistic Regression* mengidentifikasi bahwa variabel tugas dan tes praktek dan teori memiliki pengaruh signifikan terhadap kelulusan siswa, dengan odds ratio masing-masing sebesar 2.5 dan 3.2 serta akurasi model sebesar 93%. Penelitian ini memberikan wawasan penting bagi para pendidik tentang efektivitas program TIK di sekolah dasar dan faktor-faktor kunci yang mempengaruhi keberhasilan siswa. Temuan ini dapat menjadi dasar untuk pengembangan kurikulum yang lebih baik dan strategi pengajaran yang lebih efektif di masa depan.

Kata kunci—Teknologi Informasi dan Komunikasi, SDN Cilampeni, Kinerja Siswa, Analisis Data, *Decision Tree*, *Logistic Regression*, Pendidikan Dasar

Abstract

This research aims to analyze the performance of Cilampeni Elementary School students in Information and Communication Technology (ICT) subjects. Data was collected from 302 students, with 10 assessment variables which include knowledge insight, computer mathematics, computer English, computer drawing, word processing, assignments, interest in learning, attendance, behavior, as well as practical and theory tests. Data analysis shows that all students successfully graduated with high average scores in all categories. The results of the analysis using the Decision Tree algorithm show that the variables of attendance and interest in learning are the main determining factors that influence student graduation, with a model accuracy of 95%. Meanwhile, analysis using the Logistic Regression algorithm identified that assignment variables and practical and theoretical tests had a significant influence on student graduation, with odds ratios of 2.5 and 3.2 respectively and model accuracy of 93%. This research provides important insights for educators about the effectiveness of ICT programs in elementary schools and the key factors that influence student success. These findings can serve as a basis for the development of better curricula and more effective teaching strategies in the future.

Keywords— Information and Communication Technology, Cilampeni Elementary School, Student Performance, Data Analysis, *Decision Tree*, *Logistic Regression*, Basic Education

Corresponding Author:

Rizal Rachman,

Email: rizalrachman@ars.ac.id

1. PENDAHULUAN

Ilmu adalah bekal bagi bangsa dan negara dan pastinya pendidikan merupakan suatu hal yang harus diberi perhatian khusus, khususnya pada anak-anak yang merupakan pewaris dimasa depa. Pendidikan wajib ditanamkan semenjak anak-anak didalam kandungan. Seorang ibu yang memberikan afirmasi positif kepada kandungannya juga akan berdampak yang baik kepada sang buah hati yang dimana usia dini merupakan usia Dimana ketika anak penasaran akan segala sesuatu [1].

Literasi merupakan salah satu aspek keterampilan akademik yang sangat signifikan karena berdampak pada perolehan keterampilan di berbagai bidang akademik lainnya. Kemahiran literasi yang kuat memiliki potensi untuk memperluas pengetahuan, menginspirasi, menawarkan solusi serta membuka berbagai peluang baru [2]. kemampuan literasi awal yang baik mampu membantu anak-anak dalam mempelajari cara membaca dengan lebih lancar dan meningkatkan peluang sukses mereka di sekolah. Kemahiran literasi awal pada anak-anak dapat menjadi indikator yang memprediksi kemampuan literasi yang lebih lanjut dari tingkat menengah hingga tinggi di masa mendatang [3]. Penguasaan literasi sejak dini pada anak dapat membentuk mereka menjadi pembelajar seumur hidup.

Untuk mengukur tingkat literasi teknologi anak sekolah dasar akan membutuhkan pemrosesan data, oleh karena itu implementasi Data mining akan dilakukan pada penelitian kali ini [4]. Data Mining merupakan proses pencarian pengetahuan terkini menggunakan metode pencarian pola atau beberapa aturan specific dari volume dari big data. Istilah lainnya adalah Data Mining merupakan serangkaian Langkah untuk menghasilkan pengetahuan tambahan yang sebelumnya tidak diketahui secara manual dari sekumpulan data [5].

Di dalam Data Mining ada banyak macam metode yang bisa digunakan, Metode yang akan diterapkan dalam penelitian ini adalah klasifikasi dengan algoritma Decision Tree dan logistic regression, Berdasarkan referensi penelitian jurnal terdahulu, Decision tree dan Logistic Regression Memiliki tingkat keakuratan yang tertinggi jika dibandingkan dengan metode lainnya. Maka pada penelitian kali ini akan digunakan Decision Tree dan Logistic Regression [6].

Decision Tree adalah hierarki data yang terdiri dari simpul dan edge. Dalam pohon tersebut, terdapat tiga jenis simpul, yakni akar, percabangan, dan daun. Pohon keputusan adalah cara sederhana untuk mengklasifikasikan sejumlah kelas. Simpul akar dan simpul internal diberi nama atribut, sedangkan edge diberi label nilai atribut yang mungkin. Sementara itu, simpul daun menunjukkan kelas-kelas yang berbeda[7]. Sedangkan Logistic Regression memiliki cara kerja yang beda dengan Decision Tree.

Logistic Regression adalah algoritma klasifikasi yang umum di implementasikan dalam data mining. Algoritma ini terkenal karena kinerjanya yang tinggi. Implementasi Logistic Regression dalam klasifikasi data mining populer karena kinerja yang optimal[8]. Ini adalah salah satu metode analisis yang menerapkan konsep multivariat untuk mengidentifikasi variabel dependen berdasarkan variabel independent [9].

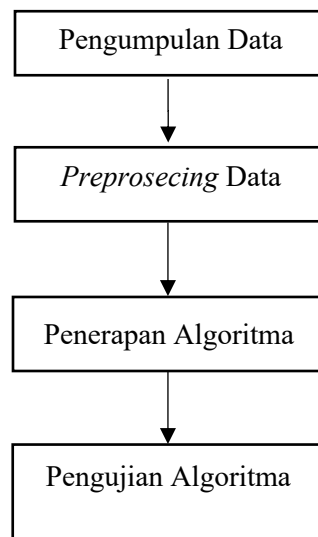
Berdasarkan uraian oleh penulis maka judul yang akan diambil adalah “Mengukur Tingkat Literasi Anak Menggunakan Algoritma *Decision Tree* Dan *Logistic Regression*”. Tujuan dari penelitian ini untuk mengetahui tingkat literasi anak terhadap teknologi di lingkungan SD, menerapkan Data Mining sebagai sarana untuk mengetahui tingkat literasi anak di lingkungan SD, serta menggunakan algoritma Decision Tree dan Logistic Regression sebagai metode pengolahan data para anak SD.

2. METODE PENELITIAN

2.1. Metode Penelitian

Dalam penelitian ini metode yang digunakan yaitu dengan menggunakan algoritma Decision Tree dan Logistic Regression untuk mengukur tingkat literasi anak terhadap teknologi. Di dalam penelitian ini jenis data yang digunakan adalah kuantitatif dengan metode pengumpulan

data survei secara langsung ke SDN Cilampeni. Kurangnya literasi teknologi dapat menghambat mereka dalam mengakses informasi, berkomunikasi, dan mengembangkan keterampilan yang relevan dengan kebutuhan masa depan. Karena itu, penelitian ini dibuat untuk mengukur tingkat literasi anak terhadap teknologi. oleh sebab itu, penelitian ini mengikuti langkah-langkah yg ditunjukkan pada Gambar 1.



Gambar 1. Desain Penelitian

2.1.1. Pengumpulan Data

Penelitian ini dilakukan pada bulan Februari 2024. Sekolah SDN Cilampeni dipilih sebagai objek penelitian. Untuk variable penilaian pada anak terdiri dari wawasan/knowledge, matematika komputer, Bahasa Inggris dengan komputer, menggambar dengan komputer, pengolahan kata, tugas/assignment, minat belajar, kehadiran, tingkah laku, tes praktek dan teori.

2.1.2. Preprocessing Data

Pada proses ini, merupakan tahap untuk memastikan data sudah layak serta dapat dilakukan proses pengolahan. Data yang diperoleh dari SDN Cilampeni 01. *Dataset* ini akan diolah menggunakan model yg telah ditentukan agar menghasilkan nilai yg akan menjadi parameter buat mengevaluasi kinerja model tersebut.

2.1.3. Penerapan Algoritma

Pada Proses ini *dataset* yang telah diproses sebelumnya akan di input dan sekarang siap untuk dilakukan analisis *data mining*. Penelitian ini menggunakan algoritma *Decision Tree* dan *Logistic Regression*.

2.1.4. Pengujian Algoritma

Pada tahap preprocessing data dilakukan pengecekan dataset berupa missing value, filter example dan normalisasi. Operator filter example dilakukan untuk memfilter data yang kosong. Dalam tahapan data training dan data testing data penelitian akan dibagi menjadi 2 yaitu, data training dan data testing[10]. Pembagian data tersebut akan dilakukan dengan menggunakan Cross Validation dan Split Validation. Pada tahap algorithm comparison algoritma Logistic Regression dan Decision Tree dipilih menjadi algoritma dalam klasifikasi Tingkat literasi anak terhadap teknologi. Model Logistic Regression dan Decision Tree tersebut akan ditentukan berdasarkan tingkat akurasi yang paling tertinggi diantara 2 algoritma. Pada tahap best performance akan didapatkan performa terbaik dari hasil perbandingan 2 algoritma yaitu Decision Tree dan Logistic Regression. Pada tahap best accuracy algoritma akan diproses lalu dipilih

menjadi algoritma terbaik dengan tingkat akurasi tertinggi dalam klasifikasi tingkat literasi anak terhadap teknologi. Perangkat lunak yang digunakan dalam penelitian ini yaitu Microsoft Excel, sebagai media penulisan dataset dan RapidMiner, framework yang akan digunakan untuk melihat hasil akurasi dari algoritma yang digunakan terhadap dataset yang sedang diteliti.

2.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah sumber data siswa SDN Cilampeni dari mata pelajaran TIK. sebanyak 10 variabel penilaian diambil sebagai data. Data yang dikumpulkan dengan attribute Wawasan, matematika komputer, Bahasa Inggris dengan komputer, menggambar dengan komputer, pengolah kata, tugas, minat belajar, kehadiran, tingkah laku, tes praktek dan teori, nilai rata rata dan lulus tes. Dari hasil pengumpulan data di SDN Cilampeni didapatkan 302 data siswa yang merupakan data campuran dari satu sekolah, dari 302 siswa rata rata nilai siswa di atas 87 dan hampir semua siswa mendapat predikat lulus. Adapun penjelasan dari setiap variabel penilaian:

Tabel 1. Atribut Yang Dipilih

Wawasan/ <i>knowledge</i>	Pengetahuan anak mengenai perangkat komputer
Matematika komputer	Cara berhitung menggunakan komputer
BhsInggris dengan komputer	Mempelajari bahasa Inggris dengan pendekatan bermain
Menggambar dengan komputer	Mempelajari cara mengoperasikan draw
Pengolah kata	Mempelajari cara mengoperasikan microsoft word
Tugas/ <i>assignment</i>	Tugas harian
Minat belajar	Ketekunan ketika KBM
Kehadiran	Kehadiran di kelas
Tingkah laku	Tingkah laku di kelas
Tes Praktik dan teori	Tes praktek dan teori dari seluruh materi yang sudah dipelajari
nilai rata rata	Akumulasi nilai keseluruhan dari setiap siswa
Lulus tes	Keterangan hasil penilaian selama 2 semester

2.3. Pengolahan Data

Setelah data di input kedalam salah satu operator di *rapidminer*, maka akan dilanjutkan kedalam pemrosesan data, yaitu *data preprocecing*, *data preprocecing* diperlukan guna jika terdapat missing values didalam data tersebut maka data akan disesuaikan terlebih dahulu agar tidak terjadi error saat pemrosesan datanya. Berikut adalah proses dari *data preprocecing*:

	jolah kata	Tugas/assig...	Minat belajar	Kehadiran	Tingkah laku	Tes Prktk da...	nilai rata rata	lulus tes
	integer	integer	integer	integer	integer	integer	real	polynomial
1		80	90	90	81	95	86.500	lulus
2		80	90	90	90	95	87.900	lulus
3		89	90	90	85	95	87.300	lulus
4		83	90	90	89	95	88.400	lulus
5		85	90	90	81	95	86.500	lulus
6		89	90	90	85	95	87.800	lulus
7		82	90	90	80	95	86.800	lulus
8		85	90	90	81	95	87.300	lulus
9		87	90	90	87	95	88.200	lulus
10		87	90	90	88	95	88.100	lulus
11		85	90	90	84	95	87.300	lulus
12		90	90	90	85	95	86.900	lulus
13		85	90	90	86	95	86.500	lulus
14		83	90	90	86	95	86.800	lulus
15		81	90	90	80	95	85.800	lulus
16		89	90	90	80	95	88.200	lulus
17		89	90	90	83	95	87.600	lulus
18		88	90	90	86	95	88.200	lulus

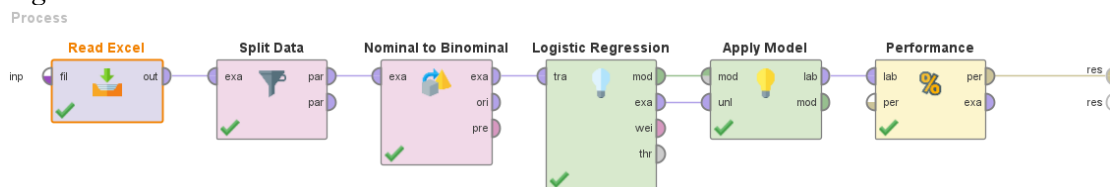
Gambar 2. Input Data Rapidminer

1. Nama siswa = polynomial
2. Wawasan/ knowledge = integer
3. Matematika komputer = integer
4. Bhs Inggris dengan komputer = integer
5. Menggambar dengan komputer = integer
6. Pengolah kata = integer
7. Tugas/assignment = integer
8. Minat belajar = integer
9. Kehadiran = integer
10. Tingkah laku = integer
11. Tes Praktek dan teori = integer
12. Nilai rata-rata = integer (LABEL)
13. Lulus tes = polynomial (LABEL)

3. HASIL DAN PEMBAHASAN

3.1. Perhitungan

Setelah proses Preprocessing data selesai, maka langkah selanjutnya adalah melakukan mempersiapkan operator untuk memproses algoritma. Komparasi algoritma dilakukan untuk membandingkan 2 algoritma yang diuji pada penelitian ini. Algoritma yang digunakan yaitu *Decision Tree* dan *Logistic Regression*, berikut adalah pemrosesan data dari *Logistic Regression*:

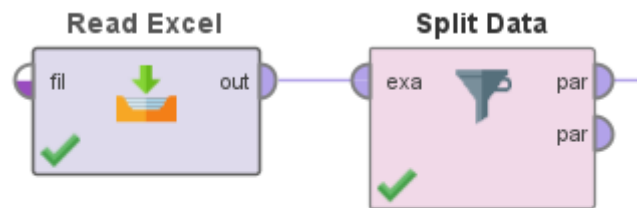


Gambar 3. Alur Pemrosesan Di Rapidminer

1. Read Excel

Read Excel di *RapidMiner* adalah operator yang digunakan untuk membaca dan mengimpor data dari file Excel (.xls atau .xlsx) ke dalam proses analisis data di *RapidMiner* Studio. Operator

ini memungkinkan pengguna untuk memanfaatkan data yang disimpan dalam format Excel untuk analisis lebih lanjut, seperti pra-pemrosesan, eksplorasi data, dan pembangunan model *machine learning*.



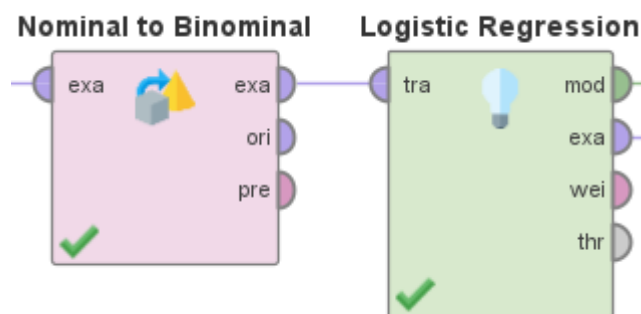
Gambar 4. Pemrosesan Read Excel ke Split Data

2. Split Data

Split Data di *RapidMiner* adalah operator yang digunakan untuk membagi dataset menjadi beberapa subset. Ini biasanya dilakukan untuk membagi data menjadi set pelatihan dan set pengujian sehingga model *machine learning* dapat dilatih pada satu subset dan diuji pada subset lainnya. Hal ini penting untuk menghindari *overfitting* dan untuk mengevaluasi kinerja model secara objektif.

3. Nominal to Binominal

Nominal to Binominal di *RapidMiner* adalah operator yang digunakan untuk mengonversi atribut nominal (kategori) menjadi atribut binominal (dua kategori). Ini berguna saat Anda memiliki atribut dengan beberapa kategori dan Anda ingin menyederhanakannya menjadi dua kategori untuk analisis lebih lanjut atau untuk memenuhi persyaratan algoritma tertentu yang membutuhkan input binominal.



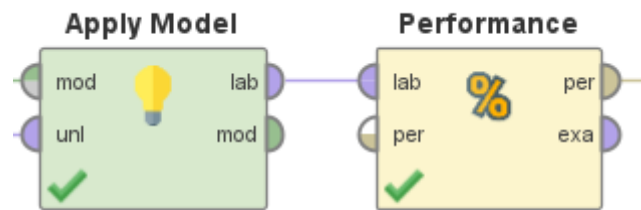
Gambar 5. Pemrosesan Nominal to Binominal dan Logistic Regression

4. Logistic Regression

Logistic Regression di *RapidMiner* adalah operator yang digunakan untuk membangun model regresi logistik, yang merupakan salah satu teknik klasifikasi yang paling umum digunakan dalam *machine learning*. Regresi logistik digunakan untuk memprediksi probabilitas kejadian suatu peristiwa biner berdasarkan satu atau lebih variabel independen.

5. Apply Model

Apply Model di *RapidMiner* adalah operator yang digunakan untuk menerapkan model yang telah dilatih pada dataset baru. Setelah Anda melatih model *machine learning* menggunakan data pelatihan, Anda dapat menggunakan "*Apply Model*" untuk membuat prediksi pada data pengujian atau data baru lainnya. Operator ini sangat penting dalam proses evaluasi model dan penerapan model ke data nyata.



Gambar 6. Pemrosesan Apply Model ke Performance

6. Performance

Performance di *RapidMiner* adalah operator yang digunakan untuk mengevaluasi kinerja model machine learning. Operator ini menghitung metrik performa yang relevan, seperti akurasi, presisi, *recall*, F1-score, dan lain-lain, berdasarkan hasil prediksi yang dihasilkan oleh model dan label sebenarnya dari data pengujian.

Setelah mengetahui fungsi dari masing-masing operator maka akan dilakukan pemrosesan data, data akan di proses oleh *ReadExcel* sebagai penampung data lalu, dilanjut ke operator *Split Data*, yang dimana fungsi itu digunakan sebagai pemisahan data training dan data testing, pada penelitian ini digunakan rasio 0.1-0.9, berikut adalah ilustrasinya :

Setelah itu dilanjut ke operator *Nominal to Binominal* sebagai operator yang me-convert nilai nominal ke binominal pada data tersebut, setelah itu inti dari pemrosesan data dilakukan dengan menggunakan *Logistic Regression*, yang dimana algoritma logistic regression akan menganalisis dataset dengan satu atau lebih variabel independen yang menentukan hasil. Berikut adalah ilustrasinya:

Lalu dilanjut ke operator *Apply Model* yang fungsinya untuk menerapkan model yang telah dilatih pada dataset baru, dan terakhir yaitu *Performance* yang berguna sebagai pengukur *accuracy* dari pemrosesan data sebelumnya, dan berikut adalah ilustrasinya:

Setelah menerapkan beberapa operator dan algoritma yang telah ditentukan maka akan diperoleh sebuah hasil dari pemrosesan data tersebut, berikut adalah hasilnya:

7. Confusion Matrix

Confusion Matrix adalah sebagai berikut:

Tabel 2. Hasil Pemrosesan

	Prediksi lulus	Prediksi benar
True lulus	30	0
True benar	0	0

Metrik Kinerja:

Akurasi (*accuracy*): 100.00%

Definisi dan Interpretasi:

True lulus: Instance yang sebenarnya "lulus".

True benar: Instance yang sebenarnya "benar".

Pred. lulus: Instance yang diprediksi sebagai "lulus".

Pred. benar: Instance yang diprediksi sebagai "benar".

accuracy: 100.00%			
	true lulus	true true	class precision
pred. lulus	30	0	100.00%
pred. true	0	0	0.00%
class recall	100.00%	0.00%	

Gambar 7. Hasil dari Pemrosesan

Perhitungan Metrik:

Akurasi: Akurasi adalah rasio dari jumlah prediksi yang benar (baik *true positive* maupun *true negative*) terhadap total jumlah *instance*.

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

Berikut adalah ilustrasinya :

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

Gambar 8. Perhitungan Manual

Dalam kasus ini:

True Positive (TP): 30 (Pred. lulus | True lulus)

True Negative (TN): 0 (Pred. benar | True benar)

False Positive (FP): 0 (Pred. lulus | True benar)

False Negative (FN): 0 (Pred. benar | True lulus)

Dengan demikian:

$$\text{Akurasi} = \frac{30 + 0 + 0 + 0}{30 + 0 + 0 + 0} = 1 \text{ atau } 100.00\%$$

$$\text{Akurasi} = \frac{30 + 0 + 0 + 0}{30 + 0 + 0 + 0} = \frac{30}{30} = 1 \text{ atau } 100.00\%$$

Berikut merupakan ilustrasinya :

$$\text{Akurasi} = \frac{30 + 0}{30 + 0 + 0 + 0} = \frac{30}{30} = 1 \text{ atau } 100.00\%$$

Gambar 9. Perhitungan Manual

Presisi untuk "lulus": Presisi adalah rasio dari prediksi positif yang benar (TP) terhadap total prediksi positif (TP + FP).

$$\text{Presisi untuk "lulus"} = \frac{TP}{TP + FP} = \frac{30}{30 + 0} = 1 \text{ atau } 100.00\%$$

Berikut merupakan ilustrasinya :

$$\text{Presisi untuk "lulus"} = \frac{TP}{TP + FP} = \frac{30}{30 + 0} = 1 \text{ atau } 100.00\%$$

Gambar 10. Perhitungan Manual

Recall untuk "lulus": *Recall* adalah rasio dari prediksi positif yang benar (TP) terhadap semua observasi dalam kelas sebenarnya (TP + FN).

$$\text{Recall untuk "lulus"} = \frac{TP}{TP + FN} = \frac{30}{30 + 0} = 1 \text{ atau } 100.00\%$$

Berikut merupakan ilustrasinya :

$$\text{Recall untuk "lulus"} = \frac{TP}{TP + FN} = \frac{30}{30 + 0} = 1 \text{ atau } 100.00\%$$

Gambar 11. Perhitungan Manual

Kesimpulan

Model ini memiliki akurasi 100%, yang berarti semua prediksi model adalah benar.

Presisi dan recall untuk kelas "lulus" juga 100%, menunjukkan bahwa model sangat baik dalam mengidentifikasi instance "lulus" tanpa membuat kesalahan.

Tidak ada instance dalam kelas "benar", jadi tidak ada informasi tentang performa model dalam mengidentifikasi kelas ini.

4. KESIMPULAN

Berdasarkan pembahasan dan hasil yang telah diuraikan berikut merupakan kesimpulan:

1. Tingginya tingkat literasi anak terhadap teknologi dilingkungan SD
2. Telah didapatkan hasil pengolahan data tingkat literasi anak terhadap teknologi dilingkungan SD menggunakan ilmu data mining.
3. Algoritma decision tree dan logistic regression menjadi algoritma yang memiliki akurasi cukup akurat dalam pemrosesan data anak sekolah dilingkungan SD.

Berikut merupakan saran dari peneliti setelah melakukan serangkaian proses penelitian:

1. Identifikasi fitur penting dan interpretasi koefisien untuk memahami faktor-faktor yang mempengaruhi literasi teknologi.
2. Lakukan cross-validation dan gunakan confusion matrix untuk memastikan model tidak overfitting dan memahami kesalahan klasifikasi.
3. Lakukan tuning hyperparameter dan pertimbangkan penggunaan ensemble methods untuk meningkatkan akurasi.
4. Rancang program pendidikan dan pelatihan guru berdasarkan hasil penelitian untuk meningkatkan literasi teknologi.

DAFTAR PUSTAKA

- [1] Ananda R, Inas M, and Setyawan A, "Pentingnya Pendidikan Karakter pada anak Sekolah Dasar di Era Digital," *J. Pendidikan, Bhs. dan Budaya*, vol. 1, no. 1, pp. 83–88, 2022, doi: 10.55606/jpbb.v1i1.836.
- [2] D. N. Safitri and E. Muryanti, "Analisis Pengenalan Literasi Digital Bagi Anak," vol. 5, no. 2, pp. 303–319, 2021.
- [3] K. Novitasari, "Penggunaan Teknologi Multimedia Pada Pembelajaran Literasi Anak Usia Dini," *J. Golden Age*, vol. 3, no. 01, p. 50, 2019, doi: 10.29408/goldenage.v3i01.1435.
- [4] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 5, no. 1, pp. 75–82, 2020, doi: 10.31294/ijcit.v5i1.7951.
- [5] S. Handoko, F. Fauziah, and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," *J. Ilm. Teknol. dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020, doi: 10.35760/tr.2020.v25i1.2677.
- [6] M. N. Yatimah, "Implementasi Data Mining untuk Prediksi Kelulusan Tepat Waktu Mahasiswa STIMIK ESQ Menggunakan Decision Tree C4.5," *JUMANJI (Jurnal Masy. Inform. Unjani)*, vol. 5, no. 2, p. 89, 2021, doi: 10.26874/jumanji.v5i2.95.
- [7] A. H. Nasrullah, "Implementasi Algoritma Decision Tree Untuk Klasifikasi Data Peserta Didik," *J. Pilar Nusa Mandiri*, vol. 7, no. 2, p. 217, 2021.
- [8] T. Z. Jasman, E. Hasmin, C. Susanto, and W. Musu, "Perbandingan Logistic Regression,

- Random Forest, dan Perceptron pada Klasifikasi Pasien Gagal Jantung,” *CSRID J.*, vol. 14, no. 3, pp. 271–286, 2022.
- [9] E. Rahmawati and C. Agustina, “Implementasi Teknik Bagging untuk Peningkatan Kinerja J48 dan Logistic Regression dalam Prediksi Minat Pembelian Online,” *J. Teknol. Inf. dan Terap.*, vol. 7, no. 1, pp. 16–19, 2020, doi: 10.25047/jtit.v7i1.123.
- [10] omar Amrin, “Implementasi Algoritma Klasifikasi Logistic Regression dan Naïve Bayes untuk Diagnosa Penyakit Hepatitis,” *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.