

ANALISIS FAKTOR YANG MEMPENGARUHI KEMAMPUAN NUMERASI SISWA SDK SURABAYA MELALUI MEDIA ULAR TANGGA: PENDEKATAN REGRESI LOGISTIK BINER DAN SVM

Ratu Bunga Prawesti Arie Salim¹, Alfisyahrina Hapsery^{2*}, Maria Trifonia Yanti Diwus³

^{1,2,3} Program Studi Statistika, Fakultas Sains dan Teknologi, Universitas PGRI Adi Buana

*e-mail: alfisyahrina@unipasby.ac.id

DOI: 10.14710/j.gauss.14.1.85-96

Article Info:

Received: 2024-09-01

Accepted: 2025-04-15

Available Online: 2025-04-29

Keywords:

Numeracy Skills; Educational Games; Binary Logistic Regression; Support Vector Machine

Abstract: *Elementary education is crucial in developing students' cognitive abilities, particularly numeracy skills essential for mathematics and daily life. Despite a curriculum designed to enhance numeracy, many students struggle with basic arithmetic concepts. Effective and engaging teaching methods, such as educational games, are needed. The Kampus Mengajar program, part of the Merdeka Belajar Kampus Merdeka (MBKM) policy, involves university students in elementary school learning, including at SDK St. Aloysius Surabaya. In this program, the snakes and ladders game serves as an educational tool to improve students' numeracy skills in a fun and interactive way. This study examines factors influencing and classifying students' numeracy abilities using Binary Logistic Regression and Support Vector Machine (SVM). The results show that binary logistic regression identifies three significant variables: gender, learning media quality, and usage frequency, with 84.06% accuracy. Meanwhile, the SVM model with an RBF kernel function achieves 92.31% accuracy. These findings highlight the importance of educational games in teaching strategies, demonstrating that SVM-based models can effectively enhance numeracy skills.*

1. PENDAHULUAN

Pendidikan dasar adalah fondasi penting untuk perkembangan kognitif dan keterampilan siswa, dengan kemampuan numerasi sebagai keterampilan mendasar yang perlu dikuasai (Handayani et al., 2022). Numerasi penting dalam matematika dan kehidupan sehari-hari karena membantu siswa memecahkan masalah (Maghfiroh et al., 2021). Meskipun kurikulum bertujuan meningkatkan kemampuan ini, banyak siswa masih kesulitan, sehingga perlu metode pengajaran yang lebih efektif (Oli et al., 2024). Penggunaan permainan edukatif, seperti ular tangga, dapat memperkenalkan konsep matematika secara menyenangkan dan interaktif, serta meningkatkan motivasi dan pemahaman siswa (Ardi & Dessty, 2023).

Program Kampus Mengajar merupakan salah satu inisiatif dalam kebijakan Merdeka Belajar Kampus Merdeka (MBKM), telah berkontribusi signifikan dalam mengembangkan metode pembelajaran inovatif di sekolah, terutama di daerah yang membutuhkan. Di SDK St. Aloysius Surabaya, program ini menggunakan permainan edukatif, seperti ular tangga, sebagai bagian dari menggunakan Program Mengajar Kampus untuk meningkatkan kemampuan numerasi siswa. Untuk mengklasifikasikan dan mengidentifikasi faktor-faktor tersebut, penelitian ini akan menggunakan pendekatan regresi logistik biner dan *Support Vector Machine* (SVM). Regresi logistik biner digunakan untuk menjelaskan keterkaitan antara variabel dependen yang bersifat biner dengan sejumlah variabel independen (Peranginangin, 2022). Dalam penelitian ini, beberapa variabel independen diantaranya kualitas materi, daya Tarik visual, frekuensi penggunaan media, dan efektivitas media.

Sementara SVM digunakan untuk meningkatkan akurasi prediksi dalam klasifikasi kemampuan numerasi siswa.

Penelitian sebelumnya menunjukkan bahwa permainan edukatif, seperti ular tangga, meningkatkan prestasi siswa dalam matematika dibandingkan metode konvensional (Hidayati, 2018; Pratiwi & Hapsari, 2020). Namun tidak ada penelitian yang mengeksplorasi secara mendalam komponen-komponen spesifik yang mempengaruhi keberhasilan tersebut dalam konteks yang lebih luas, terutama terkait dengan faktor-faktor individual siswa yang dapat mempengaruhi efektivitas permainan edukatif. Rahmawati dan Nugroho (2020) menemukan bahwa alat pendidikan interaktif, dinilai melalui regresi logistik biner, secara signifikan meningkatkan kemampuan akademik siswa. tetapi penelitian tersebut lebih fokus pada dampak alat tersebut tanpa menganalisis faktor-faktor yang dapat mempengaruhi variabel akademik melalui metode yang lebih canggih seperti regresi logistik biner dan SVM. Selain itu, Penelitian Athoillah (2018) yang menggunakan SVM dengan multi-kernel menunjukkan kemampuan klasifikasi yang baik dalam beberapa kasus, namun penerapan SVM dalam konteks pendidikan, khususnya untuk mengidentifikasi komponen yang mempengaruhi keberhasilan pembelajaran, belum banyak dieksplorasi. Penelitian ini juga belum mengaitkan hasil klasifikasi dengan desain instruksional yang lebih terperinci untuk mendukung pengajaran yang lebih efektif. Dengan mempertimbangkan latar belakang yang ada, penelitian ini bertujuan untuk menemukan komponen yang mempengaruhi dan mengklasifikasikan dengan Regresi Logistik Biner dan SVM, yang diharapkan dapat membantu merancang praktik pembelajaran yang lebih inovatif di sekolah dasar dan mendukung program Kampus Mengajar dalam meningkatkan kualitas pendidikan di Indonesia.

2. TINJAUAN PUSTAKA

Uji validitas digunakan untuk bagaimana instrumen pengukuran atau alat ukur (seperti kuesioner, tes, atau survei) secara akurat mengukur konsep atau variabel yang ingin diteliti. Hipotesis berikut ini digunakan untuk tujuan melakukan uji validitas:

H_0 : Validator memiliki pandangan yang konsisten terhadap instrumen tersebut

H_1 : Validator memiliki pandangan yang berbeda terhadap instrumen tersebut

$$r_{table} = df(N - 2, \alpha)$$

Statistik uji:

$$r_{xy} = \frac{N \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i}{\sqrt{\left(N \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2 \right) \left(N \sum_{i=1}^n Y_i^2 - \left(\sum_{i=1}^n Y_i \right)^2 \right)}} \quad (1)$$

Perhitungan koefisien korelasi diantara kedua variabel X dan Y , langkah pertama adalah mengetahui jumlah total orang yang berpartisipasi dalam survei (N). Setelah itu, hitung skor keseluruhan serta jumlah skor kuadrat untuk setiap item pernyataan ke- i dari variabel X dan Y . Kekuatan hubungan antara variabel X dan Y dapat ditentukan dengan nilai koefisien korelasi, yang didasarkan pada skor total serta jumlah skor kuadrat kedua variabel tersebut. Setelah diketahui bahwa data tersebut akurat, maka dilakukan uji reliabilitas. Menurut Sugiyono (2013), reliabilitas adalah pengukuran keterpercayaan suatu alat ukur. Suatu instrumen dianggap dapat dipercaya jika menghasilkan temuan yang konsisten setelah digunakan beberapa kali. Arikunto (2010) menjelaskan bahwa penggunaan *Cronbach's Alpha* digunakan untuk menentukan reliabilitas.

$$r_{11} = \left[\frac{K}{(K-1)} \right] \left[1 - \frac{\sum_{s=1}^n \sigma_s^2}{\sigma_t^2} \right] \quad (2)$$

Koefisien *reliabilitas* instrumen (r_{11}) dihitung dengan jumlah total item pernyataan (K), total varian item pernyataan ($\sum_{s=1}^n \sigma_s^2$), dan varian skor pernyataan total (σ_t^2).

Konsep SVM secara sederhana adalah mencari *hyperplane* optimal yang berfungsi memisahkan dua kelas atau kategori dalam sebuah ruang input. *Hyperplane* ini ditentukan oleh parameter w dan b (3) (Hastie, 2009).

$$\langle \mathbf{w}, \mathbf{x} \rangle + b = 0 \quad (3)$$

Sebuah himpunan vektor dianggap terpisah secara optimal oleh sebuah *hyperplane* jika pemisahan dapat dilakukan tanpa kesalahan, serta jarak antara vektor terdekat ke *hyperplane* mencapai nilai maksimum. *Hyperplane* dalam bentuk kanonik harus memenuhi sejumlah batasan tertentu agar pemisahan tersebut dapat tercapai pada persamaan (4).

$$y_i [\mathbf{w}, \mathbf{x}_i + b] \geq 1, i = 1, 2, \dots, l \quad (4)$$

Hyperplane optimal diperoleh dengan memaksimalkan margin $p(\mathbf{w}, b) = \frac{2}{\mathbf{w}}$ dan

memperkecil $\varphi(\mathbf{w}) = \frac{1}{2} \mathbf{w}^2$. Penyelesaian masalah ini dapat dilakukan menggunakan

berbagai teknik komputasi, salah satunya adalah metode *Lagrange Multiplier* (5).

$$L(w, b, \alpha) = \frac{1}{2} \mathbf{w}^2 - \sum_{i=1}^l \alpha_i [y_i (\mathbf{w}, \mathbf{x}_i + b) - 1] \quad (5)$$

α_i dalam persamaan (5), berperan sebagai pengganda dalam fungsi *Lagrange*. Nilai optimalnya dapat diperoleh dengan mengubah persamaan tersebut ke dalam *dual space* (6).

$$\hat{\alpha} = \arg \min_{\alpha} \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i, \mathbf{x}_j - \sum_{k=1}^l \alpha_k \quad (6)$$

Pada kenyataannya, sebagian besar permasalahan di dunia nyata jarang bersifat *linear separable*, melainkan lebih sering bersifat *non-linear separable* yaitu dalam prediksi kemampuan numerasi siswa. Faktor-faktor seperti jenis kelamin, kualitas materi, dan frekuensi penggunaan media saling berinteraksi secara kompleks dan tidak linier. Hubungan yang tidak linier ini sulit dipisahkan dengan model linear biasa. Oleh karena itu, penggunaan SVM dengan kernel RBF menjadi tepat karena mampu memetakan data ke ruang berdimensi lebih tinggi sehingga klasifikasi menjadi lebih akurat. Untuk mengatasi tantangan ini, metode SVM dimodifikasi dengan menggunakan fungsi kernel (Hastie et al., 2006) Dalam situasi *non-linear*, *hyperplane* yang bertujuan untuk memisahkan data secara optimal dilengkapi dengan variabel slack (t_i) yang perlu diminimalkan.

$$\varphi(\mathbf{w}) = \frac{1}{2} \mathbf{w}^2 + C \sum_{i=1}^l t_i \quad (7)$$

Variabel *slack* (t_i) Diterapkan untuk menangani pelanggaran terhadap pembatas $y_i [\mathbf{w}, \mathbf{x}_i + b] + t_i \geq 1$ dengan menerapkan penalti pada data yang tidak memenuhi pembatas tersebut. Penalti ini diterapkan menggunakan konstanta ongkos C . Penyelesaian optimasi (7) dapat dilakukan melalui teknik komputasi, seperti metode *Lagrange Multiplier*.

$$L(\mathbf{w}, b, \alpha, t_i, \beta) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l t_i - \sum_{i=1}^l \alpha_i [y_i (\mathbf{w}, \mathbf{x}_i + b) - 1 + t_i] - \sum_{j=1}^l \beta_j t_j \quad (8)$$

Dari persamaan (8), simbol α_i dan β_i merupakan Penambah fungsi *lag-range*. Hasil optimal dari persamaan ini dapat diperoleh dengan mentransformasikan persamaan (8) ke dalam ruang dual pada persamaan (9).

$$\hat{\alpha} = \arg \min_{\alpha} \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{k=1}^l \alpha_k \quad (9)$$

dimana $K(x_i, x_j)$ Dikenal sebagai fungsi kernel. Beberapa fungsi kernel yang umum digunakan dalam SVM adalah sebagai berikut.

1. Linier : $K(x, x') = x'x$
2. Polynomial : $K(x, x') = (\langle x, x' \rangle + 1)^p$
3. Gaussian Radial Basis Function : $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2)$

Hasil klasifikasi data x dapat dihitung menggunakan persamaan (10).

$$D(\mathbf{x}) = \sum_{i=1}^{ns} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}_j) + b \quad (10)$$

Regresi logistik, sebagaimana dijelaskan oleh Hosmer dan Lemeshow (2013), adalah metode untuk menentukan Pengaruh variabel independen terhadap variabel dependen, khususnya ketika variabel dependen tersebut bersifat biner atau memiliki dua kategori: 0 dan 1. Variabel ini didistribusikan menurut distribusi Bernoulli. Model regresi logistik biner dapat dibangun untuk beberapa variabel independen dengan menggunakan fungsi *logit* yang menetapkan hubungan antara kemungkinan terjadinya dan komponen-komponen ini. Dalam penelitian ini, model regresi logistik biner dikembangkan dengan menerapkan metode *Estimasi Maksimum Likelihood* (MLE)

$$\ln L(\beta) = \sum_{i=1}^n [y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)] \quad (11)$$

Oleh karena itu, model regresi logistik biner yang dapat dibentuk adalah sebagai berikut.

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}} \quad (12)$$

Dengan demikian, model logit untuk regresi logistik biner dapat dinyatakan sebagai berikut:

$$g(x) = \ln \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (13)$$

Model *logit* tersebut memiliki fungsi *linier* di dalam parameternya, berada dalam interval $-\infty$ sampai $+\infty$ bergantung variabel X dan bersifat *kontinyu*.

Tujuan dari uji independensi adalah untuk memastikan ada tidaknya korelasi antara dua unsur yang diselidiki, yaitu variabel independen dan variabel dependen. Arsyad (2017) Uji *Chi-Square* adalah metode statistik yang diterapkan untuk tujuan melakukan uji inferensi. Hipotesis berikut digunakan: Hipotesis awal yaitu Variabel dependen dan independen tidak saling berhubungan. sedangkan hipotesis alternatif yaitu terdapat hubungan anantara kedua variabel. Berikut Statistik uji yang diterapkan.

$$\chi_{hit}^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (14)$$

Pada uji *Chi-Square*, nilai observasi pada baris ke- i dan kolom ke- j (O_{ij}) dibandingkan dengan nilai yang diharapkan pada baris dan kolom yang sama (E_{ij}). Daerah kritis penolakan H_0 apabila $\chi_{hit}^2 > \chi_{(\alpha)(a-1)(b-1)}$. Uji signifikansi parameter dilakukan untuk mengevaluasi seberapa signifikan koefisien-koefisien dalam model yang telah dibangun sebelumnya. Proses ini melibatkan dua jenis pengujian, yaitu uji parsial dan uji simultan.

Penilaian pengaruh variabel dependen dan variabel independen secara bersamaan dilakukan dengan menggunakan teknik yang dikenal sebagai pengujian simultan. Pengujian ini didasarkan pada hipotesis sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0; j = 1, 2, \dots, p$$

Statistik uji (*Likelihood Ratio Test* (G)), yang dirumuskan sebagai berikut (Agresti, 2019):

$$G = -2 \ln \frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\sum_{i=1}^n \hat{\pi}^{y_i} (1 - \hat{\pi})^{(1-y_i)}} \quad (15)$$

Banyaknya individu yang termasuk dalam kategori 0 (n_0) dan kategori 1 (n_1) digunakan untuk menghitung proporsi masing-masing kategori dalam total sampel. Jika diterapkan tingkat signifikansi sebesar α maka keputusan tolak H_0 apabila $G^2 > X^2_{[\alpha;p]}$.

Pengujian parsial adalah suatu uji yang digunakan untuk mengevaluasi dampak substansial yang dimiliki oleh setiap variabel independen terhadap variabel yang sedang diuji. Hipotesis yang diuji adalah:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0; j = 1, 2, \dots, p$$

Statistik uji yang diterapkan adalah *Uji Wald* (W), yang dirumuskan sebagai berikut:

$$W = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)}. \quad (16)$$

Estimasi koefisien parameter ($\hat{\beta}_j$) diukur bersama dengan *standar error* (se) untuk menilai keakuratan estimasi tersebut. Jika diterapkan taraf signifikansi sebesar α maka keputusan tolak H_0 apabila $|W| > Z_{\alpha/2}$.

Tujuan dari pengujian kesesuaian model, seperti yang dinyatakan oleh Hosmer dan Lemeshow (2013), adalah untuk mengetahui apakah model yang dibangun telah sesuai dengan logistik multivariat. Pengujian ini dilakukan atas dasar hipotesis yang disajikan di bawah ini:

H_0 : Model disebut sesuai jika tidak terdapat perbedaan antara hasil pengamatan dan hasil prediksi.

H_1 : Model dikatakan tidak sesuai jika terdapat perbedaan antara hasil pengamatan dan hasil prediksi.

Statistik uji:

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n_k \pi_k)^2}{n_k \pi_k (1 - \pi_k)} \quad (17)$$

Pengamatan pada kelompok ke-k (o_k) dibandingkan dengan rata-rata taksiran peluang kelompok ke-k (π_k). Dengan g menunjukkan jumlah kelompok yang dianalisis dan n_k menyatakan banyak observasi pada kelompok ke-k.

Klasifikasi pada regresi logistic untuk mengukur apakah data yang diklasifikasikan sudah tepat (Agresti, 2019). Evaluasi ini bertujuan untuk mengidentifikasi kesalahan klasifikasi oleh *classifier*. Hasil klasifikasi dapat dievaluasi dengan tabel *confusion matrix*, yang menunjukkan jumlah data yang terklasifikasikan salah. Berikut tabel *confusion matrix* untuk klasifikasi biner.

Tabel 1. *Confusion Matrix*

Kelas Aktual	Kelas Prediksi		
	<i>Positive</i>	<i>Negative</i>	
<i>Positive</i>	<i>True Positive</i> (TP)	<i>False Positive</i> (FP)	Sensitivitas $\frac{TP}{TP + FN} \times 100\%$
<i>Negative</i>	<i>False Negative</i> (FN)	<i>True Negative</i> (TN)	Spesifisitas $\frac{TN}{TN + FP} \times 100\%$
	Akurasi $\frac{TP + TN}{TP + FN + FN + TN} \times 100\%$	Presisi $\frac{TP}{TP + FP} \times 100\%$	

Ketika mengevaluasi model klasifikasi, TP (True Positive) adalah jumlah prediksi yang benar untuk kategori positif, TN (True Negative) adalah jumlah prediksi yang benar untuk kategori negatif. FP (*False Positive*) jumlah prediksi yang tidak akurat dalam kategori positif, dan FN (*False Negative*) jumlah prediksi yang tidak akurat dalam kategori negatif.

Untuk mempermudah interpretasi model, digunakan nilai *odds ratio* (hosmer dan Lemeshow, 2013). *Odds ratio* merupakan probabilitas terkait kemungkinan suatu kejadian akan terjadi yang dinyatakan dengan pernyataan kuantitatif. Menurut (Cahyani et al., 2021), *odds ratio* memiliki beberapa sifat. Jika $OR = 1$, peluangnya sama di kedua grup. $OR > 1$ berarti grup pertama memiliki peluang lebih tinggi, sementara $OR < 1$ berarti grup pertama memiliki peluang lebih rendah. OR harus ≥ 0 dan mendekati 0 jika *odds* grup pertama mendekati 0, sedangkan OR mendekati positif tak terhingga jika *odds* grup kedua mendekati 0. Adapun persamaan *odds ratio* hal tersebut dapat dirumuskan sebagai berikut:

$$\psi = \frac{\pi(1)/1 - \pi(1)}{\pi(0)/1 - \pi(0)} \quad (18)$$

Media berperan sebagai perantara yang menghubungkan pengirim dengan penerima pesan. Dalam pembelajaran, media yang menarik dapat meningkatkan keterlibatan dan aktivitas siswa. Media pendidikan membantu menciptakan interaksi efektif antara guru, media, dan siswa untuk mengatasi kendala dalam proses belajar. Contoh media pembelajaran berbasis permainan adalah ular tangga, yang dikenal luas dan memudahkan siswa dalam penggunaannya, serta meningkatkan antusiasme mereka dalam belajar (Arsyad, 2017).

3. METODE PENELITIAN

Penelitian ini memanfaatkan data primer yang diperoleh melalui penyebaran kuesioner kepada 69 siswa kelas IV dan V di SDK St. Aloysius Surabaya. Kuesioner berisi pertanyaan tentang sikap siswa terhadap permainan ular tangga sebagai media pembelajaran kreatif untuk mengembangkan kemampuan berhitung. Sampel dalam penelitian ini dipilih menggunakan metode *purposive sampling*, yaitu teknik pemilihan sampel yang didasarkan pada penilaian subjektif peneliti (Sugiyono, 2010). Penelitian ini menilai dampak variabel terhadap kemampuan numerasi siswa dengan menggunakan permainan ular tangga sebagai alat bantu pembelajaran. Variabel dependen dan independen yang digunakan tercantum di Tabel 2.

Tabel 2. Variabel Penelitian Media Pembelajaran Ular Tangga

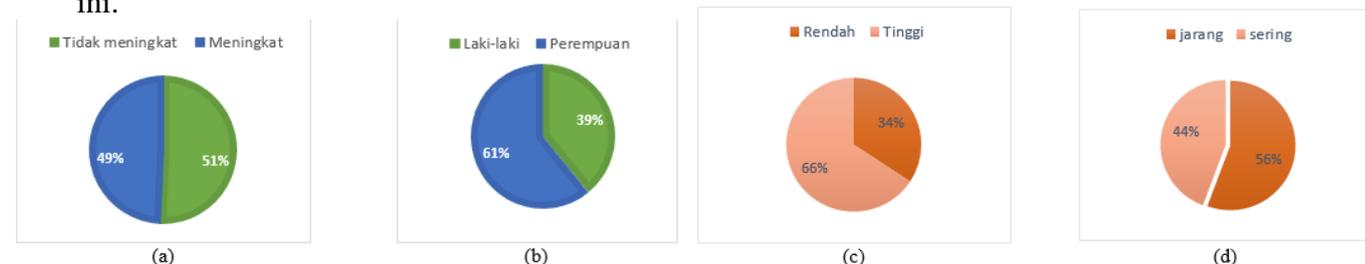
Variabel	Keterangan	Variabel	Keterangan
Usia (X ₁)	0 : Kurang dari 11 tahun 1 : Lebih dari sama dengan 11 tahun	Daya Tarik Visual (X ₆)	0 : Kurang Menarik 1 : Menarik
Jenis Kelamin (X ₂)	0 : Laki-Laki 1 : Perempuan	Efektivitas Media (X ₇)	0 : Kurang Efektif 1 : Efektif
Pekerjaan Orang tua (X ₃)	0 : Wiraswasta 1 : Karyawann Swasta	Frekuensi Penggunaan Media (X ₈)	0 : Jarang 1 : Sering
Kelas (X ₄)	0 : Kelas 4 1 : Kelas 5	Kemampuan Numerasi Siswa (Y)	0 : Tidak meningkat 1 : Meningkatkan
Kualitas Materi (X ₅)	0 : Kualitas Rendah 1 : kualitas Tinggi	Daya Tarik Visual (X ₆)	0 : Kurang Menarik 1 : Menarik

Langkah-langkah yang ditempuh dalam melakukan penelitian Dengan memanfaatkan regresi logistik biner dan Support Vector Machine (SVM). meliputi serangkaian proses sistematis yang melibatkan beberapa tahapan sebagai berikut:

1. Melakukan pengujian validitas dan reliabilitas pada data yang berskala.
2. Menjelaskan data dengan menggunakan statistik deskriptif untuk memahami karakteristik variabel.
3. Melakukan klasifikasi menggunakan regresi logistik biner melalui langkah-langkah:
 - a. Menggunakan uji independensi *Chi-Square* Untuk menganalisis hubungan antara variabel dependen (Y) dan variabel independen (X).
 - b. Melaksanakan uji signifikansi secara simultan menggunakan uji G (*Likelihood Ratio*)
 - c. Menguji signifikansi secara terpisah dengan uji *Wald*.
 - d. Menerapkan metode *Backward Elimination* untuk menghapus variabel yang tidak berpengaruh.
 - e. Menggunakan Hosmer-Lemeshow Goodness of Fit Test untuk menilai tingkat kesesuaian model.
 - f. Menghitung dan menginterpretasi ketepatan klasifikasi model.
 - g. Menginterpretasi *odds ratio*.
4. Melakukan klasifikasi menggunakan *Support Vector Machine* (SVM) melalui langkah-langkah berikut:
 - a. Membagi data menjadi dua kelompok, yaitu data pelatihan (80%) dan data pengujian (20%).
 - b. Menetapkan bobot parameter pada SVM dengan menggunakan fungsi kernel RBF.
 - c. Membuat model SVM dengan menerapkan fungsi kernel linier dan RBF.
 - d. Mengklasifikasikan data testing berdasarkan model SVM yang telah dibangun.
 - e. Menghitung akurasi klasifikasi model
5. Menarik kesimpulan.

4. HASIL DAN PEMBAHASAN

Langkah awal dalam penelitian ini adalah melakukan uji validitas untuk memastikan setiap indikator instrumen mengukur objek secara tepat, seperti Kualitas Materi (X_5), Daya Tarik Visual (X_6), Efektivitas Media (X_7) dan Frekuensi Penggunaan Media (X_8), yang semuanya terbukti valid dengan $p\text{-value} < 0,05$. Selanjutnya, dilakukan uji reliabilitas yang membuktikan bahwa seluruh variabel memiliki nilai *Cronbach Alpha* di atas 0,6, sehingga data dinyatakan konsisten. Karakteristik responden dan faktor-faktor yang mempengaruhi kemampuan numerasi siswa dapat diamati melalui statistik deskriptif yang disajikan berikut ini.



Gambar 1. Persentase Siswa berdasarkan (a) Kemampuan Numerasi Siswa, (b) Jenis Kelamin, (c) Kualitas Materi, (d) Frekuensi Penggunaan Media

Gambar 1 menunjukkan bahwa media pembelajaran ular tangga berdampak positif pada peningkatan numerasi siswa, dengan 51% siswa mengalami peningkatan. Dari sisi jenis kelamin, 61% laki-laki dan 39% perempuan, menunjukkan relevansi media ini untuk semua gender. Sebanyak 65% siswa menilai kualitas materi sebagai "tinggi," sementara 35% menilainya "rendah." Dari segi penggunaan, 55% siswa jarang menggunakan media ini, sedangkan 45% rutin. Data ini mencerminkan bahwa media ini dapat diterapkan secara efektif. Uji independensi akan dilakukan untuk menganalisis hubungan antara kemampuan numerasi dan faktor-faktor lain, dengan hasil disajikan pada Tabel 3.

Tabel 3 menunjukkan variabel Usia (X_1) dan Kelas (X_4) memiliki $p\text{-value}$ lebih besar dari 0,1, yang berarti faktor usia dan kelas tidak memiliki hubungan dengan kemampuan numerasi siswa. Sehingga, kedua faktor tersebut tidak dimasukkan dalam model. Langkah berikutnya adalah melakukan uji signifikansi parameter, baik secara simultan maupun individu, untuk mengidentifikasi faktor-faktor yang mempengaruhi kemampuan numerasi siswa.

Tabel 3. Hasil Uji Independensi

Variabel	df	Chi-Square	P-value	Keterangan
Usia (X_1)	1	0,134	0,714	Tidak Terdapat Hubungan
Jenis Kelamin (X_2)	1	14,827	0,000	Terdapat Hubungan
Pekerjaan Orang tua (X_3)	1	14,678	0,000	Terdapat Hubungan
Kelas (X_4)	1	0,129	0,719	Tidak Terdapat Hubungan
Kualitas Materi (X_5)	1	20,359	0,000	Terdapat Hubungan
Daya Tarik Visual (X_6)	1	22,808	0,000	Terdapat Hubungan
Efektivitas Media (X_7)	1	12,514	0,000	Terdapat Hubungan
Frekuensi Penggunaan Media (X_8)	1	22,28	0,000	Terdapat Hubungan

Uji secara simultan secara teori pada Sub bab 2,3 point a dan b sesuai dengan hipotesis. Pengujian dilakukan untuk mengetahui apakah faktor-faktor yang diuji mempengaruhi kemampuan numerasi siswa secara simultan. Didapatkan nilai $\chi^2_{(0,1;8)}$ sebesar 13,362 dan

nilai G^2 sebesar 47,582. Karena $G^2 > \chi^2_{(0,01;8)}$, maka hipotesis nol ditolak. Dengan demikian, hal ini mengindikasikan bahwa terdapat setidaknya satu faktor yang berpengaruh terhadap kemampuan numerasi siswa. Setelah pengujian ini, dilakukan uji lebih lanjut untuk melihat pengaruh setiap faktor secara terpisah, dan hasilnya disajikan pada Tabel 4.

Tabel 4. Uji Signifikansi Parsial Kemampuan Numerasi Siswa SDK

Variabel	Nilai Estimasi	Standard Error	Wald	P-value	Keputusan
(Intercept)	-20,293	2662,796	-0,008	0,994	Tidak Signifikan
$X_{2(1)}$	0,650	0,849	0,766	0,444	Tidak Signifikan
$X_{3(1)}$	18,020	2662,796	0,007	0,995	Tidak Signifikan
$X_{5(1)}$	0,922	1,252	0,737	0,461	Tidak Signifikan
$X_{6(1)}$	1,561	1,055	1,480	0,139	Tidak Signifikan
$X_{7(1)}$	-0,875	0,960	-0,912	0,362	Tidak Signifikan
$X_{8(1)}$	2,280	0,871	2,618	0,009	Signifikan

Tabel 4 menunjukkan bahwa satu-satunya variabel yang berpengaruh terhadap Kemampuan Numerasi Siswa (Y) adalah Frekuensi Penggunaan Media (X_8), yang memiliki p -value di bawah 0,1. Model yang terbentuk dari Tabel 4 adalah

$$\hat{g} = -20,29 + 0,65X_{2(1)} + 18,02X_{3(1)} + 0,922X_{5(1)} + 1,561X_{6(1)} - 0,875X_{7(1)} + 2,38X_{8(1)}$$

Hasil uji parsial tidak dapat membantu menemukan faktor yang mempengaruhi kemampuan numerasi, sehingga dilakukan pemilihan model terbaik dengan metode *backward elimination* dengan hasil seleksi disajikan pada Tabel 5.

Tabel 5. Seleksi *Backward Elimination*

Variabel	Nilai Estimasi	Standard Error	Wald	P-value	Keputusan
(Intercept)	-3,236	0,906	-3,571	0,000	Signifikan
$X_{2(1)}$	1,320	0,746	1,769	0,077	Signifikan
$X_{5(1)}$	2,038	0,903	2,255	0,024	Signifikan
$X_{8(1)}$	2,018	0,695	2,903	0,004	Signifikan

Tabel 5 menunjukkan variabel Jenis Kelamin (X_2), Kualitas Materi (X_5), dan Frekuensi Penggunaan Media (X_8) memiliki p -value kurang dari 0,1, yang menunjukkan bahwa faktor jenis kelamin, kualitas materi pembelajaran, dan frekuensi penggunaan media pembelajaran berpengaruh terhadap kemampuan numerasi siswa di SDK St. Aloysius Surabaya. Mengacu pada estimasi parameter yang diperoleh, persamaan regresi logistik biner yang terbentuk adalah sebagai berikut:

$$\pi(x) = \frac{\exp(-3,236 + 1,320X_{2(1)} + 2,038X_{5(1)} + 2,018X_{8(1)})}{1 + \exp(-3,236 + 1,320X_{2(1)} + 2,038X_{5(1)} + 2,018X_{8(1)})} \quad (19)$$

Dengan demikian, model regresi *logit* biner yang dihasilkan adalah sebagai berikut.

$$\hat{g} = -3,236 + 1,320X_{2(1)} + 2,038X_{5(1)} + 2,018X_{8(1)} \quad (20)$$

Model pada persamaan (21) menunjukkan bahwa siswa perempuan, kualitas materi yang baik, dan seringnya penggunaan permainan ular tangga berkontribusi pada peningkatan kemampuan numerasi siswa. Koefisien positif pada faktor-faktor ini mengindikasikan bahwa setiap faktor tersebut berperan dalam meningkatkan kemampuan numerasi. Model ini merupakan fungsi linier dari transformasi *logit*, yang mempermudah estimasi parameter regresi. Fungsi peluang dapat dihitung dengan persamaan berikut.

$$\pi(x) = \frac{\exp(-3,236 + 1,320(1) + 2,038(1) + 2,018(1))}{1 + \exp(-3,236 + 1,320(1) + 2,038(1) + 2,018(1))} = 0,894731, \text{ dengan } 1 - \pi(x) = 0,105 \quad (21)$$

Fungsi peluang menunjukkan bahwa probabilitas peningkatan kemampuan numerasi berdasarkan ketiga faktor tersebut adalah 0,895, sementara probabilitas tidak meningkat adalah 0,105. Hal ini menunjukkan adanya peluang besar untuk peningkatan kemampuan

numerasi siswa. Langkah selanjutnya adalah melakukan uji kesesuaian model guna memastikan bahwa model yang digunakan dapat menggambarkan data dengan baik.

Uji kesesuaian model menunjukkan bahwa nilai χ^2 yang diperoleh sebesar 1,533, yang lebih kecil dari nilai $\chi^2_{(0,1;2)}$ sebesar 4,605 dan memiliki *p-value* sebesar 0,465 yang melebihi 0,1. Ini mengindikasikan bahwa model yang digunakan sudah sesuai, karena tidak ada perbedaan antara hasil yang diamati dan hasil yang diprediksi. Selanjutnya, langkah berikutnya adalah menganalisis sejauh mana model ini tepat dalam mengklasifikasikan data, yang disajikan dalam Tabel 6.

Tabel 6. Tabel Klasifikasi

Aktual	Prediksi		Total
	Tidak Meningkatkan	Meningkat	
Tidak Meningkatkan	26	9	35
Meningkat	2	32	34
Total	28	41	69
Akurasi	Sensitivitas	Spesifisitas	Presisi
0,8406	0,7429	0,9412	0,9286

Tabel 6 menunjukkan bahwa akurasi model regresi logistik biner mencapai 84,06%, yang berarti model ini mampu memprediksi dengan benar sebesar 84,06% dari seluruh data yang dianalisis. Ini menunjukkan bahwa model bekerja dengan sangat baik secara keseluruhan dalam memprediksi kemampuan numerasi siswa, baik yang menunjukkan peningkatan maupun yang tidak. Selain itu, presisi model mencapai 92,86%, yang berarti bahwa ketika model memprediksi ada siswa yang mengalami peningkatan kemampuan numerasi, 92,86% dari prediksi tersebut benar-benar akurat. Ini menunjukkan bahwa model sangat tepat dalam mengidentifikasi siswa yang benar-benar mengalami peningkatan kemampuan numerasi. Langkah selanjutnya adalah interpretasi model dengan melihat nilai *odds ratio* yang terdapat pada Tabel 7.

Tabel 7. Odds Ratio

Variabel	Koefisien	Odds Ratio
X ₂₍₁₎	1,320	3,742
X ₅₍₁₎	2,038	7,672
X ₈₍₁₎	2,018	7,519

Tabel 7 menunjukkan bahwa peluang siswa perempuan berkemampuan numerasi lebih tinggi 3,742 kali dibandingkan siswa laki-laki. Selain itu, materi berkualitas tinggi memiliki peluang meningkatkan kemampuan numerasi 7,672 kali lebih tinggi dibandingkan materi berkualitas rendah. Hal serupa juga terlihat pada siswa yang sering menggunakan permainan ular tangga, di mana mereka berpeluang meningkatkan kemampuan numerasi 7,519 kali lebih tinggi dibandingkan siswa yang jarang menggunakannya.

Setelah melakukan analisis regresi logistik biner, langkah berikutnya adalah menggunakan pendekatan *Support Vector Machine* (SVM) dengan penerapan kernel RBF. Data dibagi 80:20 untuk *training* dan *testing*. Parameter *Cost* (*C*) dan *Gamma* (γ) dioptimalkan untuk mencapai akurasi terbaik, dengan hasil parameter terbaik $C = 1$ dan $\gamma = 0,1$, menghasilkan akurasi 0,9231 pada data *training*. Fungsi kernel yang digunakan adalah *Radial Basis Function* (RBF) yang menggunakan γ sebesar 0,1, menggunakan rumus

$$K(x_i, x_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) = \exp(-0,1 \|\mathbf{x}_i - \mathbf{x}_j\|^2) \text{ dengan fungsi } \textit{hyperline}$$

$D(x) = \sum_{i=1}^{36} \alpha_i y_i = \exp(-0,1 \|\mathbf{x}_i - \mathbf{x}_j\|^2) + 0,000$. Radial Basis Function (RBF) digunakan dalam Support Vector Machine (SVM) karena kemampuannya menangani data yang tidak

terpisah secara linear. Kernel ini memetakan data ke dimensi yang lebih tinggi, memungkinkan pemisahan yang lebih baik menggunakan *hyperplane*. Selain itu, RBF bersifat fleksibel dan dapat digunakan dalam berbagai kasus tanpa perlu menentukan hubungan eksplisit antar fitur. Dengan parameter γ (gamma) yang dioptimalkan, RBF juga membantu mengontrol kompleksitas model, sehingga mengurangi risiko *overfitting*. Meskipun bekerja di dimensi tinggi, perhitungannya tetap efisien karena hanya bergantung pada jarak antar titik dalam ruang fitur. Pada kasus ini, penggunaan RBF dengan $\gamma = 0,1$ menghasilkan akurasi optimal sebesar 0,9231 pada training data, menunjukkan bahwa model mampu menangkap pola data dengan baik. Model yang dibentuk dari data pelatihan digunakan untuk mengklasifikasikan data pengujian. Siswa diklasifikasikan memiliki kemampuan berhitung yang meningkat jika nilainya lebih besar dari nol, dan tidak meningkat jika nilainya kurang dari nol. Tabel 8 menyajikan hasil klasifikasi SVM dengan kernel RBF pada data *testing*.

Tabel 8. *Confussion Matrix* SVM ($C = 1, \gamma = 0,1$)

Aktual	Prediksi		Total
	Tidak Meningkat	Meningkat	
Tidak Meningkat	6	1	7
Meningkat	0	6	6
Total	6	7	13
Akurasi	Sensitivitas	Spesifisitas	AUC
0,9231	1,0000	0,8571	0,9285

Tabel 8 menunjukkan bahwa SVM dengan kernel RBF memiliki kinerja klasifikasi yang sangat baik dalam memprediksi kemampuan numerasi siswa, dengan akurasi 92,31%. Model ini berhasil mengklasifikasikan data secara tepat dan konsisten. Nilai AUC sebesar 0,9285 menunjukkan kemampuan model dalam membedakan siswa yang mengalami peningkatan numerasi dan yang tidak. AUC yang mendekati 1 menegaskan bahwa model ini sangat efektif dan akurat dalam mengidentifikasi kedua kelompok siswa.

Penelitian terdahulu menunjukkan bahwa permainan edukatif dapat meningkatkan prestasi siswa dalam matematika, namun belum mengeksplorasi faktor-faktor spesifik yang mempengaruhi keberhasilan tersebut. Selain itu, meskipun regresi logistik biner telah digunakan untuk menilai efektivitas alat pendidikan interaktif, belum ada kajian yang menggombinasikannya dengan metode klasifikasi yang lebih kompleks seperti SVM. Sementara itu, penelitian sebelumnya tentang SVM dalam pendidikan lebih berfokus pada klasifikasi umum tanpa mengaitkannya dengan desain instruksional. Hasil penelitian ini mengisi gap tersebut dengan mengidentifikasi faktor-faktor yang mempengaruhi numerasi siswa serta menunjukkan efektivitas SVM dengan kernel RBF dalam mengklasifikasikan kemampuan numerasi secara akurat, yang dapat membantu merancang strategi pembelajaran yang lebih inovatif.

5. KESIMPULAN

Hasil analisis menunjukkan bahwa regresi logistik biner mengidentifikasi tiga variabel signifikan yang memengaruhi kemampuan numerasi siswa di SDK St. Aloysius Surabaya, yaitu Jenis Kelamin, Kualitas Materi, dan Frekuensi Penggunaan Media. Model ini memiliki akurasi 84,06% dengan sensitivitas 74,29%, spesifisitas 94,12%, dan presisi 92,86%. Sementara itu, model SVM dengan kernel RBF menunjukkan akurasi lebih tinggi, yaitu 92,31%, dengan sensitivitas 100%, spesifisitas 85,71%, dan AUC 92,85%. Kedua model menunjukkan hasil klasifikasi yang baik. Temuan ini menegaskan bahwa program MBKM Kampus Mengajar berperan penting dalam peningkatan numerasi siswa melalui interaksi langsung, serta penyediaan media dan materi yang kontekstual, sehingga mendukung proses pembelajaran yang lebih efektif dan inovatif di sekolah dasar.

DAFTAR PUSTAKA

- Agresti, A. (2019). *An Introduction to Categorical Data Analysis* (3rd ed.). Wiley
- Ardi, S. D. K., & Dessty, A. (2023). Media pembelajaran ular tangga untuk meningkatkan motivasi belajar numerasi siswa di sekolah dasar. *Buletin Pengembangan Perangkat Pembelajaran*, 5(1), 1–9.
- Aridinanti, L. (2021). Pemodelan literasi membaca siswa di daerah terpencil menggunakan regresi logistik biner. *Jurnal LeECOM (Leverage, Engagement, Empowerment of Community)*, Universitas Ciputra Surabaya.
- Arikunto, S. (2010). *Prosedur penelitian: Suatu pendekatan praktik* (Edisi revisi). Rineka Cipta.
- Ariska, Y. F., & Hapsery, A. (2022). Regresi Logistik Biner untuk Memprediksi Faktor yang Mempengaruhi Nilai Akhir Peserta Didik SMK Sepuluh Nopember Sidoarjo dalam Efektivitas Pembelajaran Pasca Pandemi. *VARIANCE: Journal of Statistics and Its Applications*, 4(2), 89-98.
- Arsyad, A. (2017). *Media Pembelajaran*. Jakarta: PT RajaGrafindo Persada.
- Athoillah, M. (2018). Klasifikasi Kendaraan Bermotor Dengan Multi Kernel Support Vector Machine. *Buana Matematika: Jurnal Ilmiah Matematika Dan Pendidikan Matematika*, 8(1), 1-8.
- Cahyani, E. R. S., Khotimah, K., Agustin, R., Sari, A. E. N., & Hapsery, A. (2021). Dampak Perubahan Perilaku Mahasiswa dalam Pembelajaran Daring. *Inferensi*, 4(2), 121-127.
- Dantes, N., & Handayani, N. N. L. (2021). Peningkatan literasi sekolah dan literasi numerasi melalui model blanded learning pada siswa kelas v sd kota singaraja. *Widyalyaya: Jurnal Ilmu Pendidikan*, 1(3), 269-283
- Hastie, T. Et. All. (2009). *Springer Series In Statistics The Elements Of Statistical Learning. The Mathematical Intelligencer*, 27(2), 83–85.
- Hastie, T., Tibshirani, R., James, G., & Witten, D. (2006). *An Introduction To Statistical Learning*, Springer Texts. In Springer Texts (Vol. 102).
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley.
- Kemendikbud. (2020). *Buku Panduan Merdeka Belajar - Kampus Merdeka*. Direktorat Jenderal Pendidikan Tinggi Kemdikbud RI.
- Maghfiroh, F. L., Amin, S. M., Ibrahim, M., & Hartatik, S. (2021). Keefektifan Pendekatan Pendidikan Matematika Realistik Indonesia Terhadap Kemampuan Literasi Numerasi Siswa Di Sekolah Dasar. *Jurnal Basicedu*, 5(5), 3342-3351.
- Nirwana, R. (2022). Penggunaan Media Permainan Edukatif Ular Tangga Untuk Meningkatkan Hasil Belajar Matematika Materi Bangun Ruang Kelas V Min 2 Mojokerto. *Jurnal Madrasah Ibtidaiyah*, 1(01), 48-64.
- Oli, M. A., Dhiu, K. D., Ngura, E. T., & Sayangan, Y. V. (2024). Penggunaan Media Papan Ular Tangga Untuk Meningkatkan Pemahaman Numerasi Bagi Siswa Kelas III Di SDK Bejo. *EDUKASIA: Jurnal Pendidikan Dan Pembelajaran*, 5(1), 691-702.
- Peranginangin, I. K. M. B. (2022). Pendugaan Parameter Regresi Logistik Biner Menggunakan Maximum Likelihood Estimation (MLE) (Studi Kasus Kabupaten/Kota Provinsi Jawa Timur).
- Sahitayakti, R. P., & Fithriasari, K. (2016). Klasifikasi Kesejahteraan Rumah Tangga Di Provinsi Papua Dengan Metode Regresi Logistik Dan Support Vector Machine. *Jurnal Sains Dan Seni ITS*, 4(2).
- Sugiyono. (2013). *Metode penelitian kuantitatif, kualitatif, dan R&D* (Edisi ke-16). Alfabeta