

## ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI TINGKAT KEPUASAN SISWA MTS SURYA BUANA MENGGUNAKAN METODE REGRESI LOGISTIK

**Moh. Saleh, Totok Chamidy, Suhartono**

Magister Informatika, UIN Maulanan Malik Ibrahim Malang  
Jalan Gajayana No.50 Malang, Indonesia  
200605210015@student.uin-malang.ac.id

### ABSTRAK

Sebagai sebuah lembaga pendidikan, sekolah memiliki tanggung jawab yang besar dalam memberikan pelayanan yang berkualitas kepada siswa-siswinya. Hal ini penting agar sekolah dapat bersaing efektif dengan berbagai sekolah lain yang terus bermunculan. Dengan kemajuan teknologi saat ini, evaluasi tingkat kepuasan belajar siswa dapat dijalankan melalui metode data mining, salah satunya menggunakan pendekatan Regresi Logistik. Metode ini bertujuan untuk mengidentifikasi faktor-faktor yang memengaruhi kepuasan siswa. Regresi Logistik bekerja dengan cara menghitung probabilitas kelas dari sebuah data. Berdasarkan hasil penelitian, dapat dibuat kesimpulan bahwa kepuasan siswa di MTs Surya Buana Malang bisa diprediksi dan dievaluasi dengan menggunakan teknik data mining yang memanfaatkan algoritma regresi logistik. Hal ini bertujuan untuk memprediksi tingkat kepuasan siswa dengan memanfaatkan algoritma regresi logistik. Hasil analisis memberikan gambaran tentang faktor-faktor yang perlu diperhatikan oleh pihak sekolah dalam meningkatkan tingkat kepuasan siswa, terutama dalam hal peningkatan lingkungan sekolah dengan p-value 0.03, serta fasilitas dan infrastruktur yang memiliki p-value yang sangat rendah yakni 0.00. Selain itu Algoritma regresi logistik dapat memiliki akurasi yang sangat baik yakni 87.69% yang tergolong puas dengan menggunakan confusion matrix.

**Kata kunci :** Data Mining, Algoritma Regresi Logistik, Kepuasan Siswa

### 1. PENDAHULUAN

Sebagai sebuah lembaga pendidikan, sekolah memiliki tanggung jawab yang besar dalam memberikan pelayanan yang berkualitas kepada siswa-siswinya. Hal ini penting agar sekolah dapat bersaing efektif dengan berbagai sekolah lain yang terus bermunculan. Oleh karena itu, situasi ini menekankan perlunya sekolah dan setiap unit atau program yang ada di dalamnya untuk berkomitmen secara kuat dalam upaya berkelanjutan untuk meningkatkan mutu pendidikan yang diselenggarakan, dengan tujuan utama memenuhi dan melampaui harapan serta kepuasan siswa dan orang tua.

Dalam rangka memberikan layanan terbaik, lembaga pendidikan harus terus meningkatkan *profesionalisme* dalam memberikan layanan pendidikan yang maksimal. Selain sarana prasarana yang memadai, juga harus didukung layanan akademik yang baik, sehingga dapat terjadi pembentukan karakter yang sesuai dengan visi misi sekolah. Hal ini bertujuan agar siswa dapat memperoleh sekolah favorit dan juga membekali semua siswanya dengan pemahaman agama yang lebih baik.

Kepuasan dan ketidakpuasan pelanggan terhadap suatu produk atau jasa memiliki dampak signifikan pada pola perilaku selanjutnya [1]. Apabila seorang pelanggan merasa puas, kemungkinan besar akan kembali membeli produk atau jasa yang sama, dan pelanggan yang puas cenderung memberikan referensi positif (*word of mouth*) kepada calon konsumen yang mereka kenal. Sebaliknya, pelanggan yang tidak puas mungkin akan mengembalikan produk, mengajukan keluhan, dan akhirnya berbagi pengalaman buruk mereka tentang organisasi tersebut kepada orang lain.

Dalam kasus ekstrem, pelanggan yang tidak puas bisa mengajukan gugatan terhadap organisasi tersebut, yang berpotensi mengakibatkan kerugian besar bagi organisasi tersebut. Selain itu, cara lain yang digunakan oleh pelanggan untuk "menghukum" sekolah yang tidak mampu memberikan kepuasan adalah dengan "pindah" ke sekolah lain yang dianggap lebih mampu memberikan kepuasan.

Kepuasan pelanggan sangat dipengaruhi oleh kualitas jasa yang sesuai dengan harapan konsumen. Oleh karena itu, sekolah sebagai penyedia layanan pendidikan seharusnya memberikan penekanan pada pemberian layanan berkualitas untuk memperoleh kepercayaan dari siswa, orang tua, dan masyarakat luas.

Dengan kemajuan teknologi saat ini, evaluasi tingkat kepuasan belajar siswa dapat dijalankan melalui metode data mining, salah satunya menggunakan pendekatan Regresi Logistik. Metode ini bertujuan untuk mengidentifikasi faktor-faktor yang memengaruhi kepuasan siswa. Regresi logistik bekerja dengan cara menghitung probabilitas kelas dari sebuah data.

Pemilihan regresi logistik untuk penelitian ini didasarkan pada kesesuaian dengan sifat data dan tujuan penelitian. Alasan utamanya adalah karena variabel utama yang dilihat (kepuasan siswa) diukur dalam kategori seperti "puas," atau "tidak puas." Regresi logistik dirancang khusus untuk mengatasi jenis data kategorikal ini. Metode ini membantu kita mengukur seberapa besar kemungkinan kepuasan atau ketidakpuasan siswa terjadi berdasarkan faktor-faktor yang akan diamati. Kelebihan lainnya adalah regresi logistik dapat menangani berbagai jenis variabel, baik

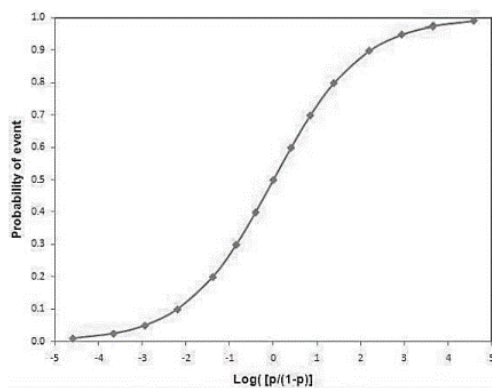
yang kategorikal maupun kontinu, dan dapat memberikan fleksibilitas dalam menganalisis data.

## 2. TINJAUAN PUSTAKA

### 2.1. Regresi Logistik

Regresi logistik adalah salah satu algoritma pembelajaran mesin yang handal digunakan untuk klasifikasi data dengan target bertipe kategori [2]. Regresi Logistik merupakan suatu metode statistik yang digunakan untuk mengevaluasi korelasi antara variabel prediktor (variabel independen) dan variabel respons (variabel dependen) yang bersifat biner atau kategorikal. Variabel respons pada regresi logistik hanya memiliki dua nilai atau kategori, seperti Ya dan Tidak, Sukses dan Gagal, atau Benar dan Salah. Metode ini cocok untuk mengatasi data yang bersifat dikotomis, yaitu data yang hanya memiliki dua kategori, dan memberikan informasi mengenai sejauh mana variabel prediktor mempengaruhi kemungkinan terjadinya suatu kejadian pada variabel respons. Analisis regresi logistik sering digunakan dalam berbagai bidang, termasuk ilmu sosial, kesehatan, ekonomi, sains, dan teknik, untuk memahami hubungan dan dampak faktor-faktor terhadap hasil tertentu [3]. Jenis regresi logistik melibatkan regresi logistik biner, regresi logistik multinomial, dan regresi logistik ordinal.

Regresi Logistik adalah jenis algoritma pembelajaran mesin klasifikasi statistik. Dasar dari algoritma ini diturunkan dari fungsi sigmoid.



Gambar 1. kurva model regresi logistik

Fungsi yang memiliki bentuk seperti Gambar 1 di atas adalah sebagai berikut:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}$$

Diketahui fungsi regresi logistik, rumus ini memperlihatkan bahwa ketika  $x \rightarrow \infty, \pi(x) \downarrow 0$  jika  $\beta < 0$  dan  $\pi(x) \uparrow 1$  jika  $\beta > 0$ . Jika  $\beta \rightarrow 0$  kurvanya cenderung membentuk garis horizontal dan jika modelnya dipenuhi dengan  $\beta = 0$  maka variabel respon biner  $Y$  akan saling bebas terhadap  $x$ .

Model regresi logistik dipakai untuk menganalisis keterkaitan antara satu variabel respon

dan beberapa variabel prediktor. Variabel respon ini berupa data kualitatif dikotomi, di mana nilai 1 menunjukkan keberadaan suatu karakteristik, sementara nilai 0 menunjukkan ketiadaannya [4]. Penerapan model regresi logistik relevan ketika variabel respon menghasilkan dua kategori dengan nilai 0 dan 1, sehingga sesuai dengan distribusi Bernoulli [5].

Regresi Logistik sangat berguna dalam masalah klasifikasi biner. Karena klasifikasi saat ini yang bersangkutan adalah tipe biner dan sangat dapat diterapkan. Kelemahan dari regresi logistik adalah rentan terhadap *underfitting* dataset yang kelasnya tidak seimbang [6], sehingga menghasilkan akurasi yang rendah.

### 2.2. Confusion matrix

*Confusion matrix* merupakan salah satu instrumen yang digunakan untuk menganalisis kemampuan suatu model algoritma klasifikasi seperti regresi logistik, dalam mengenali berbagai kombinasi data [7].

	Positive	Negative	
Positive	TP	FN	TP + FN
Negative	FP	TN	FP + TN
	TP + FP	FN + TN	

Gambar 2. parameter pada *confusion matrix*

Keterangan:

TP = True Positive

TN = True Negative

FP = False Positive

FN = False Negative

### 2.3. MSE dan RMSE

Mengukur kesalahan dalam proses klasifikasi dapat dilakukan dengan menggunakan RMSE dan MSE. *Root mean square error* (RMSE) adalah nilai akar kuadrat dari jumlah kesalahan kuadrat rata-rata yang timbul dari analisis perhitungan [8]. Semakin rendah nilai RMSE, semakin baik prediksi yang dihasilkan. Nilai RMSE yang rendah mengindikasikan bahwa model estimasi mendekati nilai sebenarnya. Sebaliknya, semakin tinggi RMSE, semakin tidak akurat nilai-nilai yang dihasilkan oleh model [9].

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{1}} \dots\dots (2)$$

*Mean squared error* (MSE) merupakan metrik yang umum digunakan untuk mengevaluasi keakuratan prediksi suatu model. MSE mengukur rata-rata dari perbedaan kuadrat antara nilai yang diprediksi oleh model dengan nilai yang diamati [10].

$$MSE = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \dots\dots (3)$$

Keterangan :

$y_1, y_2, \dots, y_n$  = nilai yang diprediksi

$y_1, y_2, \dots, y_n$  = nilai yang diamati

#### 2.4. Accuration

Akurasi merupakan prediksi true atau tingkat dari nilai kebenaran yang dihasilkan dari model algoritma dalam melakukan klasifikasi, yang mencerminkan sejauh mana prediksi model cocok dengan nilai sebenarnya [7].

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \dots (4)$$

#### 2.5. Precision

Precision adalah ukuran keakuratan suatu model algoritma yang dapat dihitung dengan membandingkan jumlah True Positive dengan total prediksi positif yang dilakukan oleh model [7].

$$precision = \frac{TP}{TP + FP} \dots (5)$$

#### 2.6. Recall

Recall adalah ukuran kelengkapan suatu model algoritma yang dapat dihitung dengan membandingkan jumlah data yang benar-benar positif dengan total data yang sebenarnya positif [7].

$$recall = \frac{TP}{TP + FN} \dots (6)$$

#### 2.7. Kualitas Pelayanan

Kualitas pelayanan (*service quality*) dapat dinilai dengan membandingkan persepsi konsumen terhadap pelayanan yang mereka terima dengan harapan mereka terhadap atribut-atribut pelayanan suatu perusahaan atau lembaga pendidikan. Ketika pelayanan yang diterima sesuai dengan harapan, maka kualitas pelayanan dianggap baik dan memuaskan. Jika pelayanan melebihi harapan konsumen, maka kualitas pelayanan dianggap sangat baik dan berkualitas tinggi. Sebaliknya, jika pelayanan yang diterima lebih rendah dari harapan, maka kualitas pelayanan dianggap buruk [11]. Pendapat Parasuraman, Zeithaml, dan Berry yang dikutip oleh Zurni menyatakan bahwa "kualitas pelayanan adalah perbandingan antara harapan konsumen terhadap pelayanan dengan pelayanan yang mereka terima." [12].

Menurut Kotler, sebagaimana dikutip dalam jurnal oleh Ester Sianturi dan Arfianti Novita Anwar, "kualitas pelayanan adalah totalitas karakteristik barang atau jasa yang mengindikasikan kemampuannya untuk memenuhi kebutuhan pelanggan, baik yang bersifat materiil maupun nonmateriil. Kualitas layanan jasa dapat diukur dengan sejauh mana keefektifannya, dan layanan jasa memiliki potensi untuk mengurangi kesenjangan antara harapan pelanggan dan layanan yang diberikan." [13].

Kualitas pelayanan dapat diidentifikasi dengan membandingkan persepsi pelanggan terhadap pelayanan yang mereka terima dengan harapan mereka

terhadap atribut-atribut pelayanan yang sebenarnya diinginkan dari suatu perusahaan. Hubungan antara produsen dan konsumen tidak terbatas pada saat pembelian, hubungan ini bersifat abadi dan meluas melewati masa kepemilikan produk hingga ke pelayanan purna jual [14].

Menurut Lewis & Booms, sebagaimana dikutip dalam Tjiptono Chandra dalam jurnal oleh Purwa Udiutomo, "kualitas pelayanan adalah ukuran sejauh mana tingkat pelayanan yang diberikan sesuai dengan harapan pelanggan." [15].

Dari definisi-definisi yang telah disebutkan di atas mengenai kualitas pelayanan, dapat disimpulkan bahwa kualitas pelayanan merujuk kepada semua tindakan yang dilakukan oleh perusahaan atau lembaga pendidikan, baik yang bersifat non-formal maupun formal, untuk memenuhi harapan dan keinginan pelanggan mereka. Pelayanan dalam konteks ini adalah layanan atau jasa yang disediakan oleh penyedia layanan dalam bentuk kemampuan, kemudahan, kecepatan, dan keramahan, yang diwujudkan melalui sikap dan perilaku penyedia layanan untuk memuaskan pelanggan.

#### 2.8. Kepuasan Siswa

Menurut Richard F. Gerson, kepuasan pelanggan adalah ketika pelanggan merasa bahwa harapan mereka telah terpenuhi atau bahkan melebihi harapan tersebut [16]. Kepuasan pelanggan terhadap suatu layanan dapat dijelaskan dengan membandingkan bagaimana pelanggan melihat layanan tersebut dengan apa yang mereka harapkan dari layanan yang diinginkan [17].

Menurut Newstrom & Davis, kepuasan pelanggan juga dapat diartikan sebagai perasaan senang atau tidak senang yang relatif [18]. Reaksi yang berbeda akan ditunjukkan oleh pelanggan setelah menggunakan layanan jasa, tergantung pada kebutuhan dan harapan mereka. Perasaan senang akan muncul jika kebutuhan dan harapan terpenuhi, sedangkan sebaliknya, ketidakpuasan akan timbul jika harapan dan kebutuhan tidak terpenuhi.

Dalam buku Handi, kepuasan pelanggan dijelaskan sebagai hasil akumulasi pengalaman konsumen atau pelanggan dalam menggunakan produk dan jasa [19]. Kepuasan pelanggan didefinisikan sebagai respons pelanggan terhadap perbedaan antara tingkat kepentingan yang mereka miliki sebelumnya dengan kinerja aktual yang mereka rasakan setelah menggunakan produk atau layanan tersebut [20].

Menurut buku yang ditulis oleh Nina, pelayanan dan kepuasan pelanggan menjadi tujuan utama dalam sebuah perusahaan karena perusahaan tidak akan dapat beroperasi tanpa pelanggan [21], dapat disimpulkan bahwa kepuasan pelanggan adalah respons perasaan seseorang terhadap pengalaman yang mereka alami sehubungan dengan harapan mereka. Dalam konteks ini, kepuasan pelanggan merujuk kepada kepuasan siswa sebagai pelanggan di lembaga pendidikan.

Kepuasan siswa terjadi ketika harapan mereka sesuai dengan pengalaman yang mereka alami di sekolah. Dengan kata lain, siswa akan merasa puas jika apa yang mereka harapkan sejalan dengan pengalaman yang mereka peroleh di sekolah tersebut.

### 3. METODE PENELITIAN

#### 3.1. Sumber Data

Bila dilihat dari sumber data yang digunakan, pengumpulan data dapat menggunakan data primer. Data primer adalah sumber data yang langsung memberikan data kepada pengumpul data (Sugiyono 2017). Data primer dalam penelitian ini yaitu kuesioner yang dibagikan ke 212 responden, yakni siswa MTs Surya Buana Malang. Kuesioner merupakan teknik pengumpulan data yang dilakukan dengan cara memberi sekumpulan pertanyaan atau pernyataan tertulis kepada responden untuk dijawab [22].

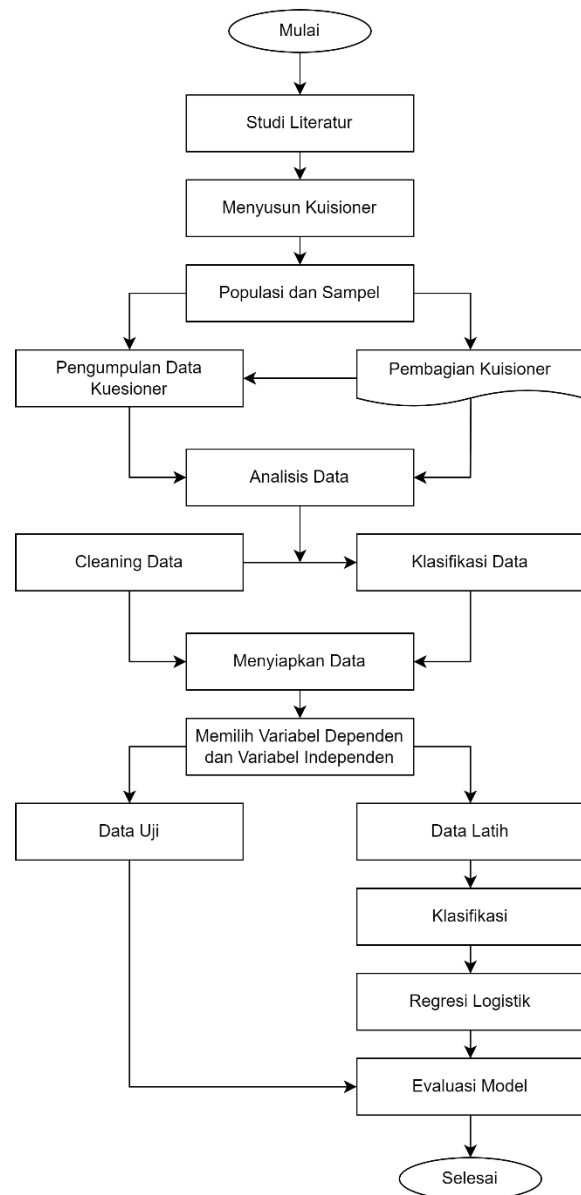
#### 3.2. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah penyebaran kuesioner. Kuesioner merupakan alat pengumpulan data yang melibatkan penyampaian beberapa pernyataan tertulis kepada responden untuk dijawab. Dalam penelitian ini, jenis kuesioner yang digunakan adalah kuesioner tertutup. Kuesioner tertutup mengandung pernyataan tertulis beserta pilihan jawaban yang telah disediakan. Model kuesioner yang digunakan dalam penelitian ini mengadopsi skala Likert. Skala likert diterapkan dengan menyajikan sejumlah pernyataan kepada responden, kemudian dijawab dengan tingkat kesesuaian yang dapat diukur melalui lima kategori. Sebelum melaksanakan penyebaran kuesioner, beberapa pengujian akan dilakukan adalah uji validitas dan uji reabilitas.

#### 3.3. Analisis Data

Analisis data ini terdiri dari beberapa proses (gambar 3). Proses awal merupakan mengumpulkan dataset, dilanjutkan dengan proses normalisasi data, validasi data, dan proses klasifikasi. Adapun algoritma klasifikasi yang digunakan Regresi Logistik.

Regresi logistik digunakan untuk memodelkan hubungan antara variabel independen dengan variabel dependen sehingga dapat mengidentifikasi faktor-faktor yang mempengaruhi tingkat kepuasan siswa MTs Surya Buana Malang. Dan selanjutnya hasil dari implementasi regresi logistik dan menarik kesimpulan serta memberikan saran dari hasil penelitian yang sudah diperoleh. Analisis menggunakan confusion matrix untuk mengevaluasi keakuratan model regresi logistik. Gambar 3 menunjukkan flowcart proses penelitian yang dilakukan.



Gambar 3. langkah-langkah proses penelitian

Dataset yang digunakan dalam penelitian yaitu data kuisisioner yang ditujukan kepada Siswa MTs Surya Buana Malang untuk mengetahui faktor-faktor apa saja yang mempengaruhi kepuasan Siswa terhadap MTs Surya Buana Malang, Dimana proses pengumpulan datanya menggunakan google form. Pengumpulan Dataset menggunakan kuisisioner berupa pertanyaan-pertanyaan yang dibentuk menjadi kelompok-kelompok sebagai berikut:

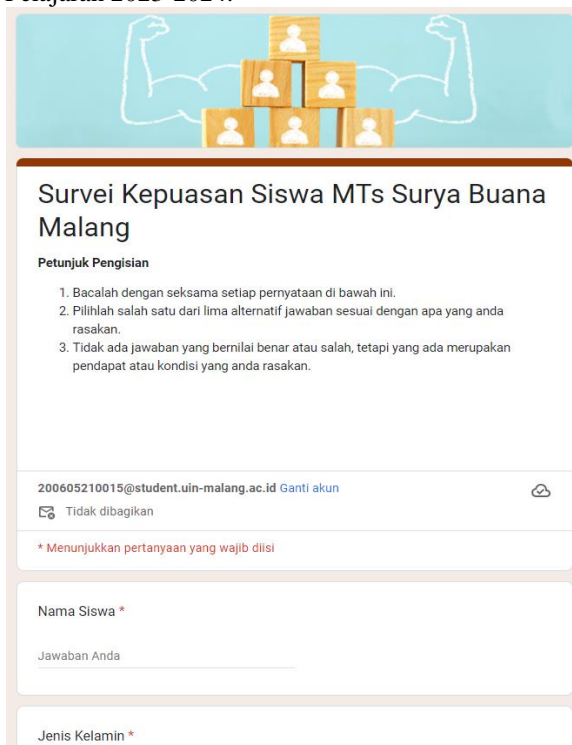
- a) Kualitas Pembelajaran
- b) Lingkungan Sekolah
- c) Hubungan Sosial
- d) Bimbingan dan Konseling
- e) Fasilitas dan Infrastruktur
- f) Pelayanan Administratif
- g) Kurikulum dan Materi Pembelajaran

Dari pengumpulan data tersebut didapatkan jumlah dataset yang sudah terkumpul sebanyak 198 data dengan 187 data yang terisi dengan benar sisa adalah data kembar dan data dengan identitas yang tidak sesuai.

#### 4. HASIL DAN PEMBAHASAN

##### 4.1. Pengolahan Data

Data yang digunakan dalam penelitian ini bersumber dari data kuesioner. Data diperoleh dari pembagian kuisisioner sebanyak 187 kuisisioner. Data yang dikumpulkan yaitu data kuisisioner yang dibagikan kepada Siswa-siswi MTs Surya Buana Malang tahun Pelajaran 2023-2024.



Gambar 4. tampilan kuisisioner

Dari hasil kuisisioner di peroleh 187 data yang terdiri dari 180 kuisisioner yang Menyatakan PUAS dan 7 kuisisioner yang Menyatakan TIDAK PUAS. Adapun pemilihan atribut data yang digunakan dapat di lihat pada tabel 4 di bawah.

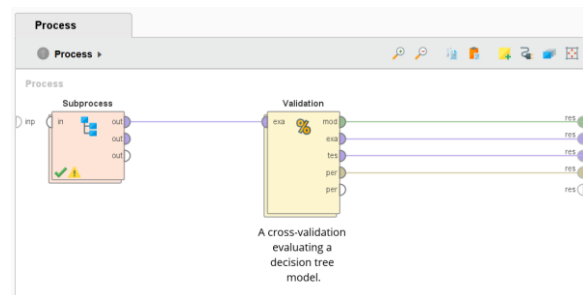
Tabel 1. Pemilihan atribut

Atribut	Detail penggunaan	
Nama	√	Id
Jenis Kelamin	X	No
Kelas	X	No
Kualitas Pembelajaran	√	Variabel bebas
Lingkungan Sekolah	√	Variabel bebas
Hubungan Sosial	√	Variabel bebas
Bimbingan dan Konseling	√	Variabel bebas
Fasilitas dan Infrastruktur	√	Variabel bebas
Pelayanan Administratif	√	Variabel bebas
Kurikulum dan Materi Pembelajaran	√	Variabel bebas
LABEL	√	Variabel terikat

Setelah data kuesioner terkumpul, selanjutnya data di olah menggunakan Microsoft Excel dengan cara mengambil rata-rata X1 (Kualitas Pembelajaran), rata-rata X2 (Lingkungan Sekolah), rata-rata X3 (Hubungan Sosial), rata-rata X4 (Bimbingan dan Konseling), rata-rata X5 (Fasilitas dan Infrastruktur), rata-rata X6 (Pelayanan Administratif), rata-rata X7 (Kurikulum dan Materi Pembelajaran), dan Y (LABEL).

##### 4.2. Pengolahan Data pada RapidMiner Studio

Pada tahap selanjutnya data akan di proses menggunakan bantuan RapidMiner Studio, untuk pengujiannya menggunakan *cross validation* yang ada di dalam RapidMiner Studio.



Gambar 5. pengolahan data rapidminer studio.

Setelah dilakukan perhitungan metode algoritma Regresi Logistik pada RapidMiner Studio, maka mendapatkan hasil seperti berikut :

**Logistic Regression Model**  
 Model Metrics Type: BinomialGLM  
 Description: N/A  
 model id: rm-h2o-model-logistic\_regression-59  
 frame id: rm-h2o-frame-logistic\_regression-59  
 MSE: 0.06767457  
 RMSE: 0.26014337  
 R<sup>2</sup>: 0.6238856  
 AUC: 0.9451685  
 pr\_auc: 0.9788435  
 logloss: 0.23747061  
 mean\_per\_class\_error: 0.15821679  
 default\_threshold: 0.45532548427581787  
 CM: Confusion Matrix (Row labels: Actual class; Column labels: Predicted class):

	Tidak Puas	Puas	Error	Rate
Tidak Puas	31	13	0.2955	13 / 44
Puas	3	140	0.0210	3 / 143
Totals	34	153	0.0856	16 / 187

Gambar 6. Hasil pengolahan algoritma regresi logistik

Hasil analisis model Regresi Logistik menggunakan metrik evaluasi seperti *Mean Squared Error* (MSE) sebesar 0.06 dan *Root Mean Squared Error* (RMSE) sebesar 0.26 menunjukkan kualitas prediksi yang baik. MSE merupakan rata-rata dari kuadrat selisih antara nilai prediksi dan nilai sebenarnya, memberikan gambaran tentang tingkat kesalahan prediksi secara keseluruhan. Nilai MSE yang rendah menandakan bahwa model memiliki tingkat kesalahan yang minim dalam memprediksi nilai target. Sedangkan RMSE merupakan akar kuadrat dari MSE, memberikan informasi tentang seberapa jauh rata-rata prediksi model dari nilai sebenarnya dalam satuan yang sama dengan variabel target. Dengan nilai RMSE yang rendah yaitu 0.26, hal

ini menegaskan bahwa model memiliki kemampuan untuk memprediksi nilai target dengan tingkat kesalahan yang rendah. Oleh karena itu, hasil analisis menunjukkan bahwa model Regresi Logistik ini dapat diandalkan dalam melakukan prediksi dengan akurasi yang tinggi.

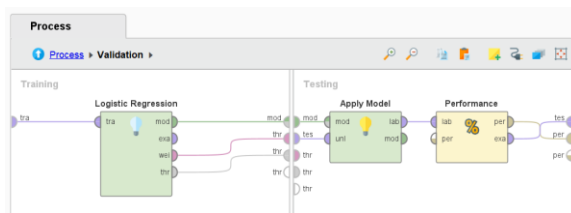
Confusion matrix (CM) menunjukkan evaluasi kinerja model dengan membandingkan kelas aktual dan kelas yang diprediksi. Terdapat dua kelas: "Tidak Puas" dan "Puas". Dari 44 sampel yang "Tidak Puas", model memprediksi 31 dengan benar dan 13 salah (error rate sebesar 0.29). Dari 143 sampel yang "Puas", model memprediksi 140 dengan benar dan 3 salah (error rate sebesar 0.02). Keseluruhan, dari 187 sampel, terdapat 16 prediksi yang salah (error rate sebesar 0.08).

Secara keseluruhan model Regresi Logistik ini memiliki kinerja yang cukup baik dengan nilai evaluasi yang menggambarkan kemampuan model dalam memprediksi kelas dengan baik. Namun perlu diperhatikan bahwa masih ada beberapa sampel yang diprediksi dengan salah, yang perlu diperbaiki dalam pengembangan model berikutnya.

### 4.3. Pengujian Analisis Regresi Logistik

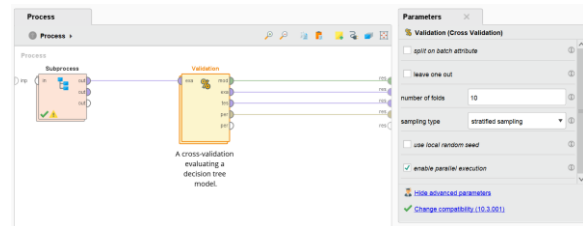
Pengujian ini dilakukan untuk mengetahui kinerja dari algoritma Regresi Logistik dalam melakukan klasifikasi terhadap kelas yang telah ditentukan dalam uji coba ini. Pengujian yang dilakukan menggunakan K-Fold Validation pada operator cross validation. K-Fold Validation ini berfungsi untuk membagi data training dan data testing pada data yang akan di uji.

Algoritma Regresi Linier melakukan training terhadap data-data yang telah dibagi oleh cross validation menjadi dua kota yaitu data training dan testing. Training sendiri terdiri dari model Regression Logistic, sedangkan untuk testing terdiri dari apply model dan performance.



Gambar 7. proses pada cross validation

Pengujian ini dilakukan dengan membagi data sebanyak 10 bagian dari 187 data kuesioner yang akan diuji. Sebanyak 10 bagian tersebut terdiri dari 9 bagian data training dan 1 bagian data testing, dalam pengujian ini akan menghasilkan nilai akurasi, precision dan recall. Berikut ini merupakan pengujian algoritma Regression Logistic pada rapidminer studio dengan menggunakan cross validation untuk pengujian model dengan number of validation 10.



Gambar 8. pengujian K-Fold Validation 10

Hasil pengujian menghasilkan accuracy, precision, dan recall sebagai berikut:

```

accuracy: 87.69% +/- 5.04% (micro average: 87.70%)
ConfusionMatrix:
True:   Tidak Puas   Puas
Tidak Puas: 27       6
Puas:    17          137

precision: 89.44% +/- 5.11% (micro average: 88.96%) (positive
class: Puas)
ConfusionMatrix:
True:   Tidak Puas   Puas
Tidak Puas: 27       6
Puas:    17          137

recall: 95.71% +/- 6.90% (micro average: 95.80%) (positive
class: Puas)
ConfusionMatrix:
True:   Tidak Puas   Puas
Tidak Puas: 27       6
Puas:    17          137
    
```

Gambar 8. Hasil pengujian menghasilkan accuracy, precision, dan recall

Data tersebut merupakan hasil evaluasi performa dari suatu model klasifikasi dengan metrik yang mencakup akurasi, presisi, dan recall. Akurasi model ini mencapai 87.69% dengan deviasi sebesar +/- 5.04%, dihitung menggunakan metode mikro rata-rata. Micro average adalah teknik penghitungan yang mengambil rata-rata dari setiap kelas dalam suatu dataset. Dari matriks kebingungan (confusion matrix), dapat dilihat bahwa model berhasil mengklasifikasikan sebagian besar instance dengan benar, dengan 27 instance dari kelas "Tidak Puas" dan 137 instance dari kelas "Puas" terklasifikasikan dengan benar. Selanjutnya, presisi model untuk kelas "Puas" adalah 89.44% dengan deviasi sebesar +/- 5.11%. Presisi adalah ukuran dari keakuratan positif yang diidentifikasi oleh model. Kemudian, recall atau "tingkat keberhasilan yang diprediksi" untuk kelas "Puas" adalah 95.71% dengan deviasi sebesar +/- 6.90%. Recall menggambarkan seberapa baik model dapat mengidentifikasi instance yang seharusnya termasuk dalam suatu kelas. Dari matriks kebingungan yang sama, model memiliki kecenderungan untuk mengklasifikasikan dengan baik kelas "Puas". Kesimpulannya, model ini menunjukkan kinerja yang baik dalam mengklasifikasikan data ke dalam kelas "Puas", meskipun masih ada ruang untuk peningkatan konsistensi dan kestabilan performa.

### 4.4. Hasil Pengujian Analisis Regresi Logistik

Hasil analisis menunjukkan koefisien untuk setiap atribut, yang mencerminkan seberapa kuat atau seberapa lemah hubungan antara atribut tersebut dengan tingkat kepuasan siswa. Analisis hasil pengujian dapat dilihat pada tabel berikut :

Tabel 2. hasil pengujian

Attribute	Coefficient	Std. Coefficient	Std. Error	z-Value	p-Value
Kualitas Pembelajaran	0.65	0.42	0.51	1.27	0.20
Lingkungan Sekolah	1.57	1.04	0.72	2.16	0.03
Hubungan Sosial	0.08	0.05	0.59	0.13	0.89
Bimbingan dan Konseling	-0.82	-0.61	0.47	-1.76	0.08
Fasilitas dan Infrastruktur	3.49	2.73	0.78	4.74	0.00
Pelayanan Administratif	0.43	0.30	0.57	0.76	0.45
Kurikulum dan Materi Pembelajaran	0.18	0.13	0.58	0.31	0.76

Data yang terlihat dalam tabel 2 merupakan hasil dari analisis regresi logistik, yang bertujuan untuk memahami hubungan antara beberapa atribut dengan probabilitas tingkat kepuasan siswa. Setiap atribut memiliki koefisien yang mengindikasikan seberapa besar dampaknya terhadap probabilitas tingkat kepuasan siswa, serta standar koefisien dan kesalahan standar yang membantu mengukur akurasi dari estimasi tersebut.

Kualitas Pembelajaran menunjukkan koefisien positif yang mengindikasikan bahwa peningkatan kualitas pembelajaran cenderung meningkatkan probabilitas keberhasilan, namun nilai p-value yang tinggi (0.20) menunjukkan bahwa hubungan ini tidak signifikan secara statistik. Lingkungan Sekolah memiliki koefisien positif yang tinggi, menandakan bahwa lingkungan sekolah yang baik secara signifikan meningkatkan probabilitas keberhasilan (p-value = 0.03).

Hubungan Sosial memiliki koefisien yang rendah dan p-value yang tinggi (0.89), menunjukkan bahwa hubungan sosial tidak memiliki pengaruh yang signifikan terhadap probabilitas keberhasilan. Bimbingan dan Konseling menunjukkan koefisien negatif, menandakan bahwa kualitas bimbingan dan konseling yang baik cenderung menurunkan probabilitas keberhasilan, namun hasil ini tidak signifikan secara statistik (p-value = 0.08).

Fasilitas dan Infrastruktur memiliki koefisien yang tinggi dan p-value yang sangat rendah (0.00), menandakan bahwa fasilitas dan infrastruktur yang baik memiliki pengaruh yang sangat signifikan terhadap probabilitas keberhasilan. Pelayanan Administratif dan Kurikulum dan Materi Pembelajaran, meskipun memiliki koefisien yang rendah, juga memiliki p-value yang tinggi, menunjukkan bahwa keduanya tidak memiliki pengaruh yang signifikan secara statistik terhadap probabilitas keberhasilan.

Dengan demikian, berdasarkan hasil regresi logistik ini, lingkungan sekolah dan fasilitas serta infrastruktur adalah faktor yang paling signifikan dalam memprediksi probabilitas keberhasilan suatu peristiwa, sementara faktor lainnya tidak memiliki pengaruh yang signifikan secara statistik pada tingkat kepercayaan yang sama.

#### 4.5. Pembahasan Analisis Regresi Logistik

Penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi tingkat kepuasan siswa MTS Surya Buana menggunakan metode regresi

logistik. Berdasarkan hasil analisis yang dilakukan dengan menggunakan metode tersebut, beberapa faktor dapat diidentifikasi sebagai berpengaruh signifikan terhadap tingkat kepuasan siswa.

Pertama, faktor Lingkungan Sekolah memiliki koefisien sebesar 1.57 dengan standar koefisien 1.04 dan p-value sebesar 0.03. Hal ini menunjukkan bahwa lingkungan sekolah memiliki pengaruh positif yang signifikan terhadap kepuasan siswa. Selanjutnya, Fasilitas dan Infrastruktur juga memiliki pengaruh yang sangat signifikan dengan koefisien 3.49, standar koefisien 2.73, dan p-value 0.00. Hal ini menandakan bahwa fasilitas dan infrastruktur yang baik dapat meningkatkan kepuasan siswa secara signifikan.

Selain itu, faktor Kualitas Pembelajaran juga menunjukkan pengaruh positif terhadap kepuasan siswa, meskipun tidak signifikan secara statistik dengan p-value sebesar 0.20. Namun, faktor Bimbingan dan Konseling memiliki koefisien negatif -0.82 dengan standar koefisien -0.61 dan p-value 0.08, menunjukkan bahwa bimbingan dan konseling yang kurang efektif dapat berdampak negatif terhadap kepuasan siswa.

Sementara itu, faktor Hubungan Sosial, Pelayanan Administratif, dan Kurikulum dan Materi Pembelajaran tidak menunjukkan pengaruh yang signifikan terhadap kepuasan siswa, karena memiliki p-value di atas 0.05.

Dengan demikian, hasil analisis ini memberikan gambaran tentang faktor-faktor yang perlu diperhatikan oleh pihak sekolah dalam meningkatkan tingkat kepuasan siswa, terutama dalam hal peningkatan lingkungan sekolah, fasilitas dan infrastruktur, serta kualitas pembelajaran.

#### 5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa kepuasan siswa di MTs Surya Buana Malang dapat diprediksi dan dievaluasi menggunakan teknik data mining dengan algoritma regresi logistik, membantu memperkirakan tingkat kepuasan siswa dengan menggunakan data training yang telah dikumpul sebelumnya. Analisis tersebut menggarisbawahi faktor-faktor yang perlu diperhatikan oleh pihak sekolah untuk meningkatkan kepuasan siswa, terutama dalam hal peningkatan lingkungan sekolah dengan p-value sebesar 0.03 yakni lebih kecil dari 0.05, serta fasilitas dan infrastruktur yang p-value 0.00. Selain itu Algoritma regresi logistik dapat memiliki akurasi yang sangat baik yakni 87.69% yang tergolong puas dengan menggunakan *confusion*

matrix, diharapkan sekolah dapat memanfaatkan data yang diperoleh untuk evaluasi dan peningkatan sistem pembelajaran agar lebih baik, serta tidak mengabaikan kepuasan siswa karena hal ini berpengaruh pada masa depan sekolah. Selain itu, sekolah perlu melakukan penelitian lebih lanjut tentang faktor-faktor yang memengaruhi kepuasan siswa, khususnya terkait dengan faktor lingkungan sekolah, serta fasilitas dan infrastruktur.

#### DAFTAR PUSTAKA

- [1] P. Kotler, *Marketing Management*. New Jersey: Prentice Hall International, Inc, 2000.
- [2] I. M. K. Karo, M. Farhan, M. Fudzee, S. Kasim, and A. A. Ramli, "Karonese Sentiment Analysis: A New Dataset and Preliminary Result," *JOIV Int. J. Informatics Vis.*, vol. 6, no. 2–2, pp. 523–530, 2022, [Online]. Available: [www.joiv.org/index.php/joiv](http://www.joiv.org/index.php/joiv)
- [3] I. Wayan, "Implementasi Logistic Regression dalam Sistem Diagnosa Penyakit Diabetes dengan KNN," *Elektron. J.*, vol. 11, no. 4, pp. 743–749, 2011.
- [4] F. Sepang, H. Komalig, and D. Hatidja, "Penerapan Regresi Logistik untuk Menentukan Faktor-Faktor yang Mempengaruhi Pemilihan Jenis Alat Kontrasepsi di Kecamatan Modayag Barat," *J. MIPA Unsrat Online*, vol. 1, no. 1, pp. 1–5, 2012.
- [5] A. Agresti, *Categorical Data Analysis*. New York: John Wiley and Sons, Inc., 1990.
- [6] I. M. K. Karo, M. F. M. Fudzee, S. Kasim, and A. A. Ramli, "Sentiment Analysis in Karonese Tweet using Machine Learning," *Indones. J. Electr. Eng. Informatics*, vol. 10, no. 1, pp. 219–231, Mar. 2022, doi: 10.52549/ijeei.v10i1.3565.
- [7] M. R. A. Nasution and M. Hayaty, "Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter," *J. Inform.*, vol. 6, no. 2, pp. 212–218, 2019.
- [8] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, and C. Lauw, "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier," *J. Nas. Teknol. Dan Sist. Inf.*, vol. 7, no. 1, pp. 8–17, 2021, doi: 10.25077/teknosi.v7i1.2021.8-17.
- [9] H. W. Herwanto, T. Widiyaningtyas, and P. Indriana, "Penerapan Algoritme Linear Regression untuk Prediksi Hasil Panen Tanaman Padi," *J. Nas. Tek. Elektro Dan Teknol. Inf.*, vol. 8, no. 4, p. 364, 2019, doi: 10.22146/jnteti.v8i4.537.
- [10] A. Yunishafira, "Determining the Appropriate Demand Forecasting Using Time Series Method: Study Case at Garment Industry in Indonesia," *KnE Soc. Sci.*, vol. 3, no. 10, p. 553, 2018, doi: 10.18502/kss.v3i10.3156.
- [11] P. Kotler and K. L. Keller, *Manajemen Pemasaran: Edisi Bahasa Indonesia*, 12th ed., vol. 1. Jakarta: PT Indeks, 2007.
- [12] Zurni, "Pengaruh Kualitas Pelayanan Terhadap Kepuasan Mahasiswa Menggunakan Perpustakaan USU," *J. Stud. Perpust. dan Inf.*, vol. 1, no. 1, p. 28, 2005.
- [13] E. Sianturi and A. N. Anwar, "Analisis Pengaruh Kualitas Pelayanan Pendidikan Terhadap Kepuasan Siswa Di Lingkungan Yayasan Pendidikan Beerseba (Studi Kasus Pada Smp Beerseba Pekanbaru)," *J. Progr. Stud. Manaj. Sekol. Tinggi Ilmu Ekon. Pelita Indones.*, vol. 3, no. 1, p. 4, 2015.
- [14] Sudaryono, *Manajemen Pemasaran Teori dan Implikasi*. Yogyakarta: CV Andi Offset, 2016.
- [15] P. Udiutomo, "Analisa Tingkat Kepuasan Siswa terhadap Layanan Program Smart Ekselensia Indonesia Tahun 2011," *J. Pendidik. Dompot Dhuafa*, vol. I, p. 5, 2011.
- [16] R. F. Gerson, *Mengukur Kepuasan Pelanggan*. Jakarta: PPM, 2001.
- [17] James, *Service Management, Operations, Strategy Information Technology*. New York: McGraw, 2014.
- [18] P. Sopiadin, *Manajemen Belajar Berbasis Kepuasan Peserta didik*. Bogor: Graha Indonesia, 2010.
- [19] H. Irawan, *10 Prinsip Kepuasan Pelanggan*. Jakarta: PT Elex Media Komputindo, 2002.
- [20] F. Rangkuti, *Measuring Customer Satisfaction: Teknik Mengukur dan Strategi Meningkatkan Kepuasan Pelanggan & Analisis Kasus PLN-JP*. Jakarta: PT. Gramedia Pustaka Utama, 2006.
- [21] N. Rahmayanty, *Manajemen Pelayanan Prima*. Yogyakarta: Graha Ilmu, 2010.
- [22] Sugiyono, *Metode Penelitian Pendidikan*. Bandung: Alfabeta, 2017.