

# Implementasi Metode Regresi Logistik Biner Pada Keakuratan Klasifikasi Jenis Musik (Studi Kasus Musik K-Pop dan Western Pada Spotify)

M. Imron

Prodi Matematika FST, UIN Sunan Ampel Surabaya, imron4838@gmail.com

Putroue Keumala Intan

Prodi Matematika FST, UIN Sunan Ampel Surabaya, putroue@uinsby.ac.id

---

**ABSTRAK**, Musik yang didengar memiliki keunikan sendiri yang dapat menjadi suatu karakteristik, dimana dengan adanya karakteristik tersebut lagu dapat dikelompokkan berdasarkan jenis lagu yang sama. Berdasarkan data dari *spotify* penyanyi dikelompokkan oleh variabel *Danceability*, *Energy*, *Acousticness*, *Tempo* dan beberapa variabel lain. Memanfaatkan variabel tersebut dilakukan penelitian bertujuan untuk melihat keakuratan klasifikasi jenis musik dimana musik yang akan di label adalah (1 = K-Pop) dan (0 = Western). Klasifikasi dilakukan dengan menggunakan regresi logistik biner yang merupakan suatu teknik klasifikasi dalam statistika untuk data dikotomi atau biner. Data penelitian di split dengan proporsi perbandingan 80:20 kemudian diolah menggunakan metode regresi logistik biner, model yang terbentuk dapat menghasilkan akurasi klasifikasi terhadap data sebesar 79,04% untuk data *training*, dan 79,69% untuk data *testing*.

**Kata Kunci:** Regresi Logistik Biner, Klasifikasi, Jenis Musik, Keakuratan.

---

## 1. PENDAHULUAN

Di era modern sekarang, perkembangan musik menjadi sangat cepat dengan memanfaatkan teknologi yang selalu *update*. Jika perkembangan musik dahulu hanya sebatas wilayah, maka dengan teknologi canggih sekarang ini dapat menyebar dengan pesat ke seluruh dunia (Rahmawati et al., 2021). Perkembangan musik tidak lepas dari seorang penyanyi, dimana karakteristik unik penyanyi terhadap sebuah lagu yang dibawakan berdampak sangat penting bagi pendengar. Improvisasi pada lagu dengan mempertahankan keunikan yang dimiliki penyanyi dapat memberikan informasi kepada pendengar lagu identitas dari penyanyi sehingga pelabelan musik pun terjadi [15].

Label musik yang disebarluaskan dari negara Eropa dan Amerika disebut musik *Western*. Musik *Western* merupakan seni musik yang didasarkan pada budaya orang-orang Eropa dan Amerika. Selanjutnya musik yang sedang *booming* dan sangat digandrungi oleh para

pendengar terutama remaja yaitu musik K-pop. Singkatan dari K-pop adalah Korean POP dimana musik ini awalnya merupakan sebuah musik daerah di Korea yaitu *trot* atau di Indonesia adalah dangdut, kemudian berkembang menjadi pop dan hip hop yang dibawakan oleh beberapa orang dalam bentuk grup [14]. Antar jenis musik memiliki karakteristik musik tersendiri dimana dapat dilihat dari beberapa variabel diantaranya *danceability*, energi, akustik, tempo lagu dan lainnya. Menggunakan karakteristik tersebut *Spotify* menempatkan lagu tersebut dalam kelompok yang sejenis. Klasifikasi atau pengelompokan yang dilakukan terhadap jenis musik ini dapat diukur keakuratannya menggunakan regresi logistik biner [9].

Klasifikasi adalah suatu metode dalam statistika yang bertujuan mengelompokkan data berdasarkan karakteristik yang dimiliki data kemudian dimasukkan dalam kelas tertentu. Salah metode klasifikasi statistika digunakan sebagai metode analisis pada penelitian ini adalah metode regresi logistik biner. Regresi logistik biner adalah suatu teknik klasifikasi dalam dalam statistika untuk data berkategori dikotomi atau biner dengan tujuan meneliti hubungan variabel bebas terhadap variabel terikat, dimana akan menghasilkan suatu nilai peluang atau probabilitas untuk tiap-tiap label musik yang digunakan sebagai dasar pengklasifikasian [3].

Penelitian yang telah dilakukan dengan memanfaatkan metode regresi logistik biner dalam klasifikasi diantaranya adalah pengklasifikasian nasabah pada koperasi simpan pinjam dan pembiayaan syariah dengan metode regresi logistik biner dimana variabel signifikan yang dijadikan model adalah usia dan besar pendapatan dengan pembagian data uji dan data latih 80%:20% menghasilkan akurasi klasifikasi

yang sebesar 78,67% [4]. Selanjutnya penelitian yang dilakukan oleh Ronny dkk, dimana tujuan penelitian adalah klasifikasi penerima KIP kuliah pada mahasiswa baru. Hasil penelitian dengan variabel yang berpengaruh pada model adalah jumlah penghasilan, status bentuk rumah, jumlah kendaraan bermotor yang dimiliki, dan jumlah mobil didapat akurasi dalam klasifikasi sebesar 88[11].

Latar masalah di atas dijadikan dasar dalam penelitian ini yang mana akan dilakukan analisis dalam melihat keakuratan klasifikasi pada jenis musik menggunakan variabel yang dianggap mempengaruhi pengklasifikasian dimana dalam pengimplementasian proses klasifikasi memanfaatkan metode Regresi Logistik Biner.

## 2. TINJAUAN PUSTAKA

### Jenis Musik K-Pop dan Western

*K-Pop* adalah perkembangan dari musik daerah menjadi musik modern yang berasal dari Korea. Musik K-pop mampu meninggalkan rasa puas pada pendengarnya terutama Amerika sehingga menyebabkan cepatnya penyebaran musik K-pop hingga dapat mendominasi pasar musik dunia. Musik K-pop yang modern sangat serasi dengan minat musik generasi sekarang sehingga peminat musik korea pop menjadi lebih digemari[9]. Masyarakat Korea sedari dulu telah menanamkan musik dalam sebagai bagian dari kegiatan harian yang terus dilakukan, efek yang terjadi adalah konsumsi terhadap suatu musik menjadi sangat tinggi dan ini juga yang menyebabkan cepatnya peredaran ekonomi di Korea pada bidang seni [7].

Musik *Western* adalah suatu istilah musik yang mendapatkan sentuhan dari negara barat terutama Amerika, Eropa, dan Inggris [8]. Perkembangan musik Western di Indonesia secara pesat terjadi dalam era presiden Soekarno pada tahun 1950-an hingga sekarang. Ciri yang terlihat jelas dalam musik Western adalah tempo yang dimiliki cukup cepat, dan pemanfaatan alat musik yang selalu modern [6].

### Regresi Logistik Biner

Regresi Logistik Biner diterapkan guna melihat hubungan variabel Y dengan data yang bersifat dikotomi atau biner dengan satu atau

lebih variabel independen. Biner atau dikotomi memiliki nilai 0 dan 1 [10]. Model Regresi Logistik Biner dibentuk dengan nilai  $P(Y|x)$  sebagai  $\pi(x)$ , dimana  $\pi(x)$  sebagai berikut [12].

$$\pi(x) = \left[ \frac{\exp(g(x))}{1 + \exp(g(x))} \right] \quad (2.1)$$

Fungsi di atas merupakan fungsi regresi non linier, agar menjadi fungsi linier dilakukan transformasi logit. berikut fungsi logit dari  $\pi(x)$  sebagai berikut [2].

$$\text{logit}(\pi(x)) = \ln \left[ \frac{\pi(x)}{1 - \pi(x)} \right] = g(x) \quad (2.2)$$

dimana  $g(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$

Keterangan:

$\pi(x)$  = Model Regresi Logistik Biner

$g(x)$  = Fungsi Logit Transformasi dari  $\pi(x)$

### Uji Signifikansi Model Serentak

Uji Serentak dilakukan guna melihat signifikansi variabel X terhadap variabel Y jika uji secara bersama-sama menggunakan nilai *Likelihood Ratio*. Adapun hipotesis yang diuji adalah sebagai berikut [10].

Hipotesis yang digunakan:

$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$  (Tidak ada pengaruh antara variabel X terhadap Y)

$H_1$ : terdapat minimal satu  $\beta_p \neq 0$  (ada pengaruh antara variabel X terhadap Y)

Rumus statistik uji:

$$G = -2 \ln \left( \frac{\text{Likelihood tanpa variabel bebas}}{\text{Likelihood dengan variabel bebas}} \right) \quad (2.3)$$

Jika taraf signifikan adalah  $\alpha$  keputusan hipotesis yang diambil tolak  $H_0$  ketika nilai dari  $G > \chi_p^2$  atau  $p\text{-value} < \alpha$  dan sebaliknya.

### Uji Signifikansi Model Individu

Uji signifikansi secara individu bertujuan melihat variabel bebas saat diuji secara parsial variabel mana saja yang masih tetap signifikan terhadap variabel terikat sehingga layak diikutsertakan dalam model menggunakan *Wald Test*. Adapun hipotesis yang diuji adalah sebagai berikut [5].

Hipotesis yang digunakan:

$H_0: \beta_p = 0$  (variabel X secara parsial tidak signifikan terhadap Y)

$H_1: \exists \beta_p \neq 0$  (variabel X signifikan terhadap Y)

Rumus statistik uji:

$$W^2 = \left( \frac{\hat{\beta}_p}{SE(\hat{\beta}_p)} \right) \quad (2.4)$$

Keterangan:

$\hat{\beta}_p$  = Estimator dari  $\beta_p$

$SE(\hat{\beta}_p)$  = Standar error

Jika taraf signifikan adalah  $\alpha$ , tolak hipotesis  $H_0$  saat nilai  $W^2 > \chi^2_{(\alpha,1)}$  atau  $p\text{-value} < \alpha$  dan sebaliknya.

### Odds Ratio

Odds ratio adalah suatu pengukuran yang digunakan untuk menginterpretasikan suatu koefisien dari variabel bebas atau dapat diartikan variabel bebas yang cenderung memiliki nilai tertentu saat diberikan  $x = 1$  dibandingkan jika diberikan  $x = 0$ . Secara umum odds ratio didapat dari sekumpulan probabilitas yang dibagi oleh probabilitas lainnya. Berikut rumus *Odds ratio* [1].

$$OR = \left[ \frac{\frac{\pi(1)}{1 - \pi(1)}}{\frac{\pi(0)}{1 - \pi(0)}} \right] \quad (2.5)$$

Keterangan

$OR$  = Odds Ratio

### Uji Kesesuaian Model

Uji Kesesuaian model dilakukan guna melihat kredibilitas dari model. Jika model sesuai berarti tidak ada perbedaan antara prediksi dan observasi begitu juga sebaliknya. Statistik uji yang cocok digunakan adalah *Hosmer-Lemeshow Test*. Adapun hipotesis yang digunakan adalah sebagai berikut [13].

Hipotesis yang digunakan:

$H_0$ : Model sesuai (tidak ada perbedaan antara prediksi dan pengamatan)

$H_1$ : Model tidak sesuai (ada perbedaan antara prediksi dan pengamatan)

Rumus statistik uji:

$$\hat{C} = \sum_{k=1}^g \left( \frac{O_k - n'_k \bar{\pi}_k}{(n'_k \bar{\pi}_k)(1 - \bar{\pi}_k)} \right) \quad (2.6)$$

Keterangan:

$O_k$  = nilai variabel Y pada grup ke- $k$

$n'_k$  = banyak observasi grup ke- $k$

$\bar{\pi}_k$  = rata-rata taksiran probabilitas grup ke- $k$

$g$  = banyak grup

Jika taraf signifikan adalah  $\alpha$  maka keputusan tolak  $H_0$  saat nilai  $\hat{C} > \chi^2_{(\alpha, g-2)}$  dengan  $g$  adalah grup atau  $p\text{-value} < \alpha$  dan sebaliknya.

### Keakuratan Klasifikasi

Hasil dari klasifikasi kemudian diukur keakuratan dalam perhitungan. Pengukuran keakuratan dilakukan dengan prosedur *APER (Apparent Error Rate)* yang ditampilkan dalam matriks konfusi. Matriks konfusi berisi jumlah data yang benar dan salah yang telah diklasifikasikan [5].

Tabel 1. Matriks Konfusi

Aktual	Prediksi	
	1	0
1	TP	FN
0	FP	TN

dengan perhitungan ketepatan akurasi sebagai berikut.

$$Akurasi = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \quad (2.7)$$

Keterangan:

TP = True Positive (jumlah kasus positif yang diklasifikasikan dengan benar oleh model)

FP = False Positive (jumlah kasus negatif yang salah diklasifikasikan sebagai positif oleh model)

FN = False Negative (jumlah kasus positif yang salah diklasifikasikan sebagai negatif oleh model)

TN = True Negative (jumlah kasus negatif yang diklasifikasikan dengan benar oleh model)

## 3. METODOLOGI

### Variabel dan Sumber Data Penelitian

Penelitian ini merupakan penelitian kuantitatif, data yang diolah bertipe data numerik. Berikut data sekunder yang digunakan dalam penelitian didapat dari *dataset kaggle* (<https://www.kaggle.com/kpopwesternsongs>).

Tabel 3.1 Sampel data

artist	Y	X1	X2	X3	X4	X5	X6	X7	X8
K-pop	1	0.65	0.77	-3.84	0.05	0.24	0.09	0.61	149.92
K-pop	1	0.44	0.85	-2.52	0.31	0.04	0.06	0.77	192.03
:	:	:	:	:	:	:	:	:	:
Western	0	0.79	0.57	-8.86	0.12	0.01	0.08	0.41	85.48
Western	0	0.81	0.44	-9.53	0.33	0.06	0.10	0.31	151.13

Dimana variabel yang digunakan seperti pada tabel dibawah ini.

Tabel 3.2 Variabel data

no	variabel	Nama Variabel	jenis data	skala
			kategori	
1	Y	Jenis Musik	0=Western 1=K-pop	Nominal
2	X1	Danceability	Numerik	Rasio
3	X2	Energi	Numerik	Rasio
4	X3	Kenyaringan Suara	Numerik	Interval
5	X4	Speechiness	Numerik	Rasio
6	X5	Akustik	Numerik	Rasio
7	X6	Liveness	Numerik	Rasio
8	X7	Valence	Numerik	Rasio </td
9	X8	Tempo	Numerik	Rasio

**Langkah-Langkah Regresi Logistik Biner**

Adapun tahapan penelitian yang dilakukan dengan menggunakan metode Regresi Logistik Biner sebagai berikut:

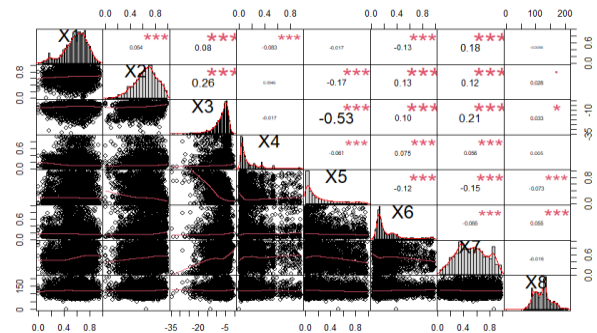
1. Melakukan identifikasi pada data dengan matriks korelasi.
2. Melakukan pengecekan kasus multikolinieritas dengan ketentuan VIF (*Variance Inflation Factor*).
3. Membagi secara random data menjadi data *training* dan data *testing* sebesar 80:20.
4. Melakukan pemodelan menggunakan metode regresi logistik biner
5. Melakukan uji serentak pada model regresi logistik biner menggunakan *Likelihood Ratio test*.
6. Melakukan uji parsial pada model regresi logistik biner, uji ini dilakukan dengan menggunakan *Wald Test*.
7. Menentukan nilai odds ratio dari model regresi logistik biner
8. Melakukan *Goodness of fit test*.
9. Melakukan perhitungan keakuratan klasifikasi model.

**4. PEMBAHASAN**

**Identifikasi Kasus Multikolinieritas**

Pemeriksaan terhadap asumsi non-multikolinieritas pada data bertujuan untuk melihat korelasi antar variabel prediktor dalam model, dimana jika dalam model terindikasi adanya kasus multikolinieritas menyebabkan parameter model tidak dapat diestimasi dengan baik sehingga model yang terbentuk menjadi bias ataupun jika teridentifikasi hasilnya tetap tervalidasi jelek. Berikut visualisasi dari matriks

korelasi data jenis musik pada Gambar 1 dibawah ini.



Gambar 4.1 Visualisasi korelasi

Dari Gambar 1 terlihat bahwa nilai korelasi pada tiap-tiap variabel independen tidak lebih dari 0,5 yang artinya kasus multikolinieritas hampir tidak terjadi. Selain itu, sebagai penguat ada atau tidaknya kasus multikolinieritas juga dapat ditentukan dengan melihat dari nilai VIF (*variance inflation factor*), berikut ditampilkan nilai VIF dari variabel data jenis musik.

Tabel 4.1 Nilai VIF

Variabel	VIF
X <sub>1</sub>	1,0696
X <sub>2</sub>	1,0363
X <sub>3</sub>	1,2500
X <sub>4</sub>	1,0374
X <sub>5</sub>	1,2205
X <sub>6</sub>	1,0779
X <sub>7</sub>	1,0690
X <sub>8</sub>	1,0064

Terlihat pada Tabel 4 dengan kriteria VIF dimana hasil nilai VIF lebih besar dari 10 maka terdapat kasus multikolinieritas pada variabel, namun dapat diamati pada tabel bahwa nilai VIF dari tiap variabel X berada di bawah nilai 10 yang menandakan variabel bebas dari kasus multikolinieritas. Dari 8 variabel yang dicek semuanya memenuhi uji asumsi non multikolinieritas yang menandakan bahwa semua variabel dapat diikutsertakan dalam pembentukan model.

**Analisis Regresi Logistik Biner**

Pemodelan diawali dengan membagi data secara random agar semua data memiliki kesempatan yang sama dalam pembentukan model. Data dibagi menjadi data *training* dan

data *testing* dengan perbandingan 80:20. Selanjutnya dilakukan uji signifikansi serentak dengan menggunakan *Likelihood Ratio Test*, dengan taraf signifikansi  $\alpha = 0,01$  nilai uji serentak ditampilkan pada Tabel 5 di bawah ini.

Tabel 4.2 Uji Serentak

Nilai statistik uji G	df	Chi-Square
1820,800	8	20,09024

Dari Tabel 5 di atas dapat dilihat bahwa nilai  $G > \chi_p^2$  yang artinya daerah keputusan tolak  $H_0$  dengan artian bahwa minimal ada satu variabel secara simultan atau serentak memiliki pengaruh terhadap variabel respon. Kemudian dilanjutkan dengan uji signifikansi secara parsial yang hasilnya ditampilkan pada **tabel 4.3** di bawah ini.

Table 4.3 Uji Parsial

Variabel	Wald	df	p-value	Chi-square
$X_1$	51,241	1	8,169E-13	
$X_2$	7,8139	1	0,005185	
$X_3$	646,726	1	2,20E-16	
$X_4$	44,4579	1	2,599E-11	
$X_5$	26,1554	1	3,15E-07	6,634897
$X_6$	179,902	1	2,20E-16	
$X_7$	7,5602	1	0,005967	
$X_8$	19,2841	1	0,00001126	

Dari Tabel 4.3 dengan taraf signifikansi  $\alpha = 0,01$  didapat nilai dari  $W^2 > \chi_{(\alpha,1)}^2 = 6,634897$  sehingga diambil keputusan tolak  $H_0$  yang berarti semua variabel berpengaruh secara signifikan walaupun diuji secara terpisah, maka tiap variabel dapat dilibatkan dalam pemodelan. Selanjutnya dibentuk model regresi logistik biner berdasarkan koefisien ( $\beta$ ) pada tabel berikut ini.

Tabel 7. Uji Estimasi Parameter

Variabel	VIF
$\beta_0$	-0,3435
$\beta_1$	1,7412
$\beta_2$	0,7047
$\beta_3$	0,5616
$\beta_4$	1,8519
$\beta_5$	1,0135
$\beta_6$	2,9528

$\beta_7$	0,5172
$\beta_8$	0,0067

Dari nilai yang berada pada Tabel 7 didapat model logit dari regresi logistik biner sebagai berikut.

$$g(x) = (-0.3435) + 1.7412x_1 + 0.7047x_2 + 0.5616x_3 + 1.8519x_4 + 1.0135x_5 + 2.9528x_6 + 0.5172x_7 + 0.0067x_8$$

Dimana berdasarkan fungsi logit yang terbentuk, dapat dibentuk fungsi peluang dari regresi logistik biner di bawah ini.

$$\pi(x) = \frac{e^{\left( \begin{matrix} (-0.3435)+1.7412x_1+0.7047x_2+ \\ 0.5616x_3+1.8519x_4+1.0135x_5+ \\ 2.9528x_6+0.5172x_7+0.0067x_8 \end{matrix} \right)}}{1 + e^{\left( \begin{matrix} (-0.3435)+1.7412x_1+0.7047x_2+ \\ 0.5616x_3+1.8519x_4+1.0135x_5+ \\ 2.9528x_6+0.5172x_7+0.0067x_8 \end{matrix} \right)}}$$

Langkah selanjutnya, sebagai interpretasi dari model dapat dipermudah dengan melihat nilai *Odds Ratio* pada tabel berikut ini.

Table 4.4 Nilai Odds Ratio

Variabel	Exp( $\beta$ )
$X_1$	5,7044
$X_2$	2,0232
$X_3$	1,7535
$X_4$	6,3719
$X_5$	2,7552
$X_6$	19,1588
$X_7$	1,6773
$X_8$	1,0067

Dari tabel nilai *Odds Ratio* di atas, model regresi logistik biner dapat ditafsirkan dimana variabel  $X_1$  berpengaruh 5,7044 kali pada jenis musik K-pop dibandingkan jenis musik Western, yang mana  $X_1$  adalah variabel *Danceability*. Selanjutnya variabel  $X_2$  lebih memiliki pengaruh 2,0232 pada K-pop dibandingkan jenis musik Western. Kemudian pada variabel  $X_3$  pengaruh pada jenis musik Western lebih rendah sebesar 1,7535 jika dibandingkan dengan jenis musik K-pop, dan selanjutnya dapat diinterpretasikan untuk tiap variabel.

### Uji Kesesuaian Model

Uji kesesuaian model atau *Goodness of fit test* dilakukan dengan melihat nilai dari *Hosmer-Lemeshow* berikut ini.

Tabel 4.5 Nilai *Hosmer-Lemeshow*

<i>Hosmer-Lemeshow test</i>	<i>df</i>	<i>Chi-square</i>	<i>p-value</i>
12,989	8	20,0902	0,1122

Model dikatakan memiliki kredibilitas yang baik secara statistik adalah ketika nilai uji *Hosmer-Lemeshow test* lebih kecil dari nilai *Chi-square* atau nilai *p-value* dari uji lebih besar dari taraf signifikan yang ditentukan dalam penelitian. Berdasarkan **Tabel 4.5** di atas dengan taraf signifikan adalah 1% maka keputusan terima  $H_0$  sebab nilai  $\hat{C} < \chi^2_{(\alpha, g-2)} = 20,0902$  atau dapat dilihat nilai  $p - value > 0.01$  yang mengindikasikan bahwa model tersebut telah sesuai dimana tidak terdapat perbedaan antara prediksi dan observasi.

### Uji Keakuratan Model

Selanjutnya hasil ketepatan klasifikasi dari jenis musik yang telah diklasifikasi ditampilkan pada tabel dibawah ini.

Tabel 4.5 Matriks Konfusi

klasifikasi	Prediksi		akurasi	error
	1	0		
<i>Training</i>				
Aktual	1	1468	444	79,04%
	0	373	1612	20,06%
<i>Testing</i>				
Aktual	1	357	108	79,69%
	0	90	420	20,31%

Dari Tabel 10 di atas terlihat bahwa dengan proporsi perbandingan 80:20 pada data kemudian diolah dengan menggunakan metode regresi logistik biner. Hasil ketepatan dari model dalam melakukan klasifikasi berdasarkan dilihat pada tabel matriks konfusi dimana saat model dilakukan klasifikasi terhadap data *training* menghasilkan akurasi sebesar 79,04%. Klasifikasi pada data *training* ini berhasil mengklasifikasi data benar pada tempat yang benar sebanyak 1468, mengklasifikasi data yang salah pada tempat yang salah sebanyak 1612, mengklasifikasi data yang seharusnya benar tetapi ditempatkan pada klasifikasi yang salah sebanyak 373 dan data yang seharusnya diklasifikasikan salah ditempatkan pada klasifikasi yang benar sebanyak 444 data.

Kemudian dari model dilakukan klasifikasi pada data *testing* menghasilkan nilai akurasi sebesar 79,69%, dimana akurasi ini berhasil mengklasifikasikan data benar pada klasifikasi benar sebanyak 357, data salah yang diklasifikasikan benar sebanyak 90, data yang salah dengan benar diklasifikasikan salah sebanyak 429 dan data benar yang diklasifikasikan salah sebanyak 108 data.

## 5. KESIMPULAN

Pada penelitian ini dengan menggunakan metode regresi logistik biner dalam melihat keakuratan pengklasifikasian jenis musik K-Pop dan Western pada aplikasi Spotify menghasilkan nilai keakuratan pada proporsi perbandingan 80:20 yaitu sebesar 79,04% pada data *training* dan 79,69% pada data *testing*. Nilai ini termasuk dalam ambang yang baik.

## 6. DAFTAR PUSTAKA

- [1] Ardianty, nayatul F., & Anggriyani, I. R. (2020). Model Dan Peluang Terjadinya Hujan Berdasarkan Kelembaban Menggunakan Regresi Logistik Biner Di Kabupaten Manokwari. *Jurnal Natura*, 16(1), 34–37.
- [2] Bera, B., Saha, S., & Bhattacharjee, S. (2020). Forest cover dynamics (1998 to 2019) and prediction of deforestation probability using binary logistic regression (BLR) model of Silabati watershed, India. *Trees, Forests and People*, 2(July). <https://doi.org/10.1016/j.tfp.2020.100034>
- [3] Fitri, R. E., Setiawan, E., Usman, M., & Aziz, D. (2022). Analisis Regresi Logistik Biner Terhadap Data Indeks Kedalaman Kemiskinan Di Indonesia Tahun 2020. *Jurnal Siger Matematika*, 03(02), 69–74.
- [4] Innassuraiya, S., Widiharih, T., & Utami, I. T. (2022). Analisis Klasifikasi Menggunakan Metode Regresi Logistik Biner Dan Bootstrap Aggregating Classification And Regression Trees (Bagging Cart) (Studi Kasus: Nasabah Koperasi Simpan Pinjam Dan Pembiayaan Syariah (Kspps)). *Jurnal Gaussian*, 11(2), 183–194. <https://doi.org/10.14710/j.gauss.v11i2.354>

58

- [5] Ndangi, W. R. A., Resmawan, R., & Djakaria, I. (2019). Perbandingan Analisis Diskriminan dan Regresi Logistik Multinomial. *Jambura Journal of Mathematics*, 1(2), 54–63. <https://doi.org/10.34312/jjom.v1i2.2100>
- [6] Pertiwi, A., & Nasution. (2014). Larangan Soekarno terhadap musik Barat tahun 1959-1967. *Journal Pendidikan Sejarah*, 2(3), 1.
- [7] Rahmawati, A., Nikmah, N. L., Perwira, R. D. A., & Rakhmawati, N. A. (2021). Analisis topik konten channel YouTube K-pop Indonesia menggunakan Latent Dirichlet Allocation. *Teknologi: Jurnal Ilmiah Sistem Informasi*, 11(1), 16–25. <https://doi.org/10.26594/teknologi.v11i1.2155>
- [8] Ramadhan, A. F., Made, D., Laksmi, S., Wiyati, W. S., Depresi, K. K., Suara, P., & Barat, T. M. (2022). “Depresi”: Sebuah Paduan Suara Akapela dengan Penggabungan Teknik Musik Barat dan Tangga Nada Selendro Betawi. *Journal Of Music*, 1(1), 19–27.
- [9] Sakinah, R. N., Hasna, S., & Wayuningsih, Y. (2022). Pengaruh Positif Fenomena K-Pop Terhadap Karakter Generasi Muda di Indonesia. *Journal on Education*, 05(01), 735–745.
- [10] Sari, E. D. N., & Irhamah, I. (2020). Analisis Sentimen Nasabah pada Layanan Perbankan Menggunakan Metode Regresi Logistik Biner, Naïve Bayes Classifier (NBC), dan Support Vector Machine (SVM). *Jurnal Sains Dan Seni ITS*, 8(2). <https://doi.org/10.12962/j23373520.v8i2.44565>
- [11] Susetyoko, R., Wiratmoko Yuwono, & Elly Purwantini. (2022). Model Klasifikasi Pada Seleksi Mahasiswa Baru Penerima KIP Kuliah Menggunakan Regresi Logistik Biner. *Jurnal Informatika Polinema*, 8(4), 31–40. <https://doi.org/10.33795/jip.v8i4.914>
- [12] Wang, J., Song, H., Fu, T., Behan, M., Jie, L., He, Y., & Shangguan, Q. (2022). Crash prediction for freeway work zones in real time: A comparison between Convolutional Neural Network and Binary Logistic Regression model. *International Journal of Transportation Science and Technology*, 11(3), 484–495. <https://doi.org/10.1016/j.ijst.2021.06.002>
- [13] Wulandari, E., & Puhadi, P. (2020). Analisis Pengaruh Faktor-Faktor Infeksi Cacing pada Balita dan Anak Umur 6-12 Tahun di Kabupaten Ende, Nusa Tenggara Timur dengan Metode Regresi Logistik Biner. *Jurnal Sains Dan Seni ITS*, 8(2). <https://doi.org/10.12962/j23373520.v8i2.43107>
- [14] Yenti, N. S., Mairiza, N., Anggraini, N., & Dkk. (2022). Dampak Budaya Korea Pop (K-Pop) Terhadap Tingkat Motivasi Belajar Mahasiswa Universitas Negeri Padang. *Jurnal Pendidikan, Bahasa, Sastra, Seni Dan Budaya*, 2, 176–191.
- [15] Zubedi, F., Aliu, M. A., Rahim, Y., & Oroh, F. A. (2021). Analisis Klasifikasi Artist Music Menggunakan Model Regresi Logistik Biner Dan Analisis Diskriminan. *Jambura Journal Of Probability And Statistics*, 2(May), 1–8.