



Prediksi Rounds Golf Dengan Algoritme Regresi Linier Berganda (Studi Kasus: Lotuslakes Golf Club)

Lala Aprianti Putri¹, Tesa Nur Padilah², Hannie³

^{1,2,3} Universitas Singaperbangsa Karawang

Abstract

Received: 02 Januari 2024

Revised : 10 Januari 2024

Accepted: 17 Januari 2024

Prediction is a process of predicting what will happen in the future. This study aims to predict golf rounds with the Multiple Regression Algorithm using the Knowledge Discovery In Database (KDD) method. The data used in this study amounted to 1863 data, with the dependent variable rounds, and 3 independent variables namely, members, guests, and golf carts. In analyzing the data, the researcher used Minitab 16 tools where the classic assumption test was carried out first which consisted of a normality test, multicollinearity test, autocorrelation test, and heteroscedasticity test and hypothesis testing. The prediction results obtained show that if the number of golf members, guests, and golf carts is large, then the predicted number of golf rounds will also increase, with an effect of 96.25%, which means that almost all of them are influenced by the member, guest, and golf cart variables. Furthermore, from the prediction results, an evaluation of the error rate was carried out, namely the value of MAD 0.038101262, MSE 1351.307149, MAPE 0.06399372, and RMSE 36.76012989. The error rate of the MAD calculation is 0.038101262, which means an error rate of 3.81% or <10% is classified as a small error or the prediction results are very accurate.

Keywords: Data mining, Prediction, Multiple Linear Regression, Golf, Rounds.

(*) Corresponding Author: Lala.aprianti18051@student.unsika.ac.id

How to Cite: Putri, L. A., Padilah, T. N., & Hannie. (2024). Prediksi Rounds Golf Dengan Algoritme Regresi Linier Berganda (Studi Kasus: Lotuslakes Golf Club). <https://doi.org/10.5281/zenodo.10521331>.

PENDAHULUAN

Olahraga merupakan suatu kegiatan fisik yang sangat bermanfaat untuk kesehatan, sekaligus wadah yang dapat mengekspresikan diri manusia. Dalam Undang-undang Nomor 3 Tahun 2005 Ayat 4 Pasal 1 Bab I tentang Sistem Keolahragaan Nasional menyebutkan olahraga merupakan segala kegiatan yang tersusun untuk mendorong, membina, dan mengembangkan potensi jasmani, rohani, dan sosial. Dalam Pasal 17 Bab IV ruang lingkup olahraga dibagi menjadi 3, yakni olahraga pendidikan, olahraga rekreasi, dan olahraga prestasi. Olahraga dapat dilakukan di tempat terbuka (*outdoor*) maupun di tempat tertutup (*indoor*). Salah satu olahraga yang dilakukan di tempat terbuka yaitu olahraga golf (Irianto, 2020).

Olahraga golf merupakan salah satu cabang olahraga *outdoor* yang dapat dilakukan secara individu maupun kelompok, yang berlomba memasukkan bola ke dalam lubang dengan jumlah pukulan paling sedikit. Asal mula olahraga golf yaitu berasal dari Skotlandia. Golf di Indonesia dikenal semenjak berdirinya klub golf batavia 1872 yakni sejak masa penjajahan Belanda dengan kondisi geografis Indonesia iklim tropis, juga alam yang bagus membuat permainan golf mulai berkembang di Indonesia dan banyak perusahaan yang berlomba menciptakan

lapangan golf yang terbaik. Salah satu perusahaan yang mendirikan lapangan golf yaitu Lotuslakes Golf Club.

Dalam 3 tahun terakhir *player* di Lotuslakes Golf Club mengalami penurunan dan jumlah *round* pun menurun juga, hal ini dikarenakan pandemi covid-19. Pada tahun 2019 ada sebanyak 38.833 *round*, tahun 2020 sebanyak 36.253 *round*, dan tahun 2021 sebanyak 34.830 *round*. Lotuslakes golf club telah menggunakan sistem *Rhapsody* sejak tahun 2010 yang merupakan sistem menangkap, memproses, dan menganalisis semua data dan transaksi dan mengintegrasikan seluruh alur operasional secara akurat. Dengan adanya sebuah sistem yang telah diterapkan di lotuslakes golf club, menimbulkan sebuah tumpukan data yang sangat banyak dan terkadang membuat sistem menjadi lambat dan error. Permasalahan yang dihadapi oleh pihak lotus lakes adalah inputan data yang hanya dibiarkan saja. Maka diperlukan suatu metode untuk mengolah data tersebut menjadi sebuah pengetahuan atau *knowledge* untuk pihak lotuslakes dalam mengambil keputusan yang bijak.

Data mining merupakan rangkaian proses pencari suatu nilai tambah dari kumpulan data berupa pengetahuan yang belum diketahui sebelumnya dan akhirnya menyimpulkannya menjadi informasi-informasi penting guna meningkatkan keuntungan, dan memperkecil pengeluaran (Baihaqi, Dianingrum & Ramadhan, 2019). Dengan menggunakan data mining, kumpulan data tersebut dapat diolah dan diprediksi untuk menghasilkan informasi yang dapat meningkatkan daya saing dan pengelolaan perusahaan.

Dalam melakukan perencanaan pihak manajemen maupun marketing lotus seringkali mengalami kesulitan. Jumlah *round* golf yang bersifat fluktuatif dan tidak dapat diperkirakan jumlah pastinya, menimbulkan perencanaan yang dibuat menjadi tidak bekerja dengan efisien. Meski tidak ada prediksi yang tingkat keberhasilannya 100% akurat, tetapi hal ini perlu diantisipasi dengan melakukan prediksi jumlah *player* yang akan bermain golf, dengan mencari metode yang dapat menghasilkan prediksi dengan nilai akurasi yang tinggi, yakni dengan pendekatan *data mining* yang dimodelkan dengan algoritme regresi linier berganda.

Metode regresi linier berganda merupakan teknik dalam menentukan hubungan antara variabel yang ingin diestimasi atau diramalkan variabel satu (variabel tak bebas) dengan variabel lain (variabel bebas) (Sella, 2020). Menurut Rizky, F., Syahra, Y., & Mariami, I. (2019) juga menjelaskan bahwa metode regresi linier berganda merupakan kajian terhadap ketergantungan antar satu variabel dengan variabel lainnya yang bertujuan membuat estimasi rata-rata nilai variabel yang telah diketahui.

Pada penelitian yang dilakukan oleh Indah Lestari Lumban Gaol, et al., (2019) yang berjudul “Implementasi Data Mining Dengan Metode Regresi Linier Berganda Untuk Memprediksi Data Persediaan Buku Pada PT. Yudhistira Ghalia Indonesia Area Sumatera Utara”, memprediksi ketersediaan stok buku berdasarkan algoritme regresi linier berganda dan menggunakan tools RapidMiner untuk pengujian persamaan regresi linearnya. Penelitian lainnya terkait regresi linier berganda juga pernah dilakukan oleh Sari Yulia Retno, (2019) terkait perbandingan analisis logika fuzzy dan regresi linier beganda dalam menuntukan produksi beras nasional yang menghasilkan nilai rata-rata kesalahan relatif dengan regresi linier berganda sebesar 0,0078 atau 0,78%, dan nilai rata-rata kesalahan relatif

peramalan dengan Logika Fuzzy sebesar 0,0930 atau 9,3%. Demikian, rata-rata kesalahan *relative* prediksi metode fuzzy lebih besar daripada metode regresi linier berganda maka dari itu regresi linear berganda lebih unggul dibandingkan dengan fuzzy. Penelitian yang berkaitan dalam memprediksi *round* golf belum ada, maka berdasarkan uraian latar belakang di atas mendorong penulis untuk membuat penelitian yang diberi judul “Prediksi *rounds* golf dengan Algoritme Regresi Linier Berganda (Studi Kasus: Lotuslakes Golf Club)”.

METODE

Metodologi yang digunakan dalam penelitian ini adalah metode *Knowledge Discovery in Database* (KDD), dengan tahapan sebagai berikut:

1. Data Selection

Pada tahap pertama dilakukan pemilihan atribut data yang diperoleh dari Lotuslakes Golf Club yang dimana ada atribut yang diperlukan dan tidak diperlukan. Data awal merupakan data laporan *player* FO lotuslakes dari 1 Januari 2017 sampai 28 Februari 2023.

2. Preprocessing/Cleaning

Pada tahap ini dilakukan pembersihan data yang tidak lengkap isinya (*missing value*) yang diatasi secara manual di Ms. Excel. Kemudian dilakukan juga pembersihan data dengan menghapuskan *outlier* yang terdapat pada dataset, yang diatasi dengan menggunakan metode boxplot. Hasil dari tahap ini adalah *dataset final* yang sudah siap untuk diaplikasikan ke algoritma.

3. Transformation

Tahap ini akan dilakukan transformasi dengan logaritma natural, yaitu dengan mentransformasikan variabel independen ke logaritma natural, dan variabel dependen tetap, sehingga data dapat masuk ke tahap selanjutnya yaitu tahap *data mining*.

4. Data Mining dengan Algoritme Regresi Linier Berganda

Tahap ini merupakan implemetasi algoritma regresi linier berganda untuk memprediksi *round* golf Lotuslakes Golf Club. Algoritma regresi linier berganda meliputi uji asumsi klasik (uji normalitas, uji multikolinieritas, uji autokorelasi, dan uji heteroskedastisitas), Pengujian Parameter (uji T dan uji F), dan koefisien determinasi.

5. Evaluation

Tahap evaluasi dilakukan dengan menghitung nilai *error* antara hasil prediksi dengan hasil sesungguhnya pada *data testing*. Perhitungan nilai *error* menggunakan *Mean Squared Error* (MSE), *Mean Absolute Deviation* (MAD), *Mean Absolute Percent* (MAPE) dan *Root Mean Square Error* (RMSE).

HASIL DAN PEMBAHASAN

HASIL

Hasil penelitian ini merupakan hasil dari penerapan metode *data mining* prediksi dengan algoritma regresi linier berganda dengan permasalahannya yaitu untuk memprediksi *round* golf berdasarkan *player member*, *guest* dan golf cart di lotuslakes golf club.


Langkah yang dilakukan pada *data selection* adalah pemilihan atribut sehingga data dapat diproses ke tahap berikutnya yaitu *preprocessing*. Adapun data didapatkan langsung dari Lotuslakes Golf Club. Data awal merupakan data rekapitulasi pemain (*player*) Member dan Guest dari 1 Januari 2017 hingga 28 Februari 2023. Dari total *records* data laporan *player* FO lotuslakes dari masing-masing atribut yaitu 2.249. Contoh data awal dapat dilihat pada Gambar 4.1.

DATE/DAY	MEMBER				GUEST				GOLFERS	ROUNDS	GOLF CART		
	W.DA : 18		W-end/NH:9		W.DAY		W. END				AM	PM	TOTAL
	AM	PM	AM	PM	AM	PM	AM	PM					
1 Sun(New Y)			81	16			49	36	182	182	39	31	70
2 Mon Open	41	4			105	41			191	200	61	26	87
3 Tue	7	2			47	6			62	76	30	4	34
4 Wed	2	1			13	8			24	24	10	5	15
5 Thu	4	2			46	15			67	79	29	9	38
6 Fri	3	2			24	6			35	45	13	5	18
7 Sat			159	25			89	131	404	404	87	85	172
8 Sun			138	16			90	77	321	321	86	49	135
9 Mon									0				0
10 Tue	2	0			65	9			76	91	34	5	39
11 Wed	2	6			42	5			55	66	24	7	31
12 Thu	2	1			31	2			36	42	19	2	21
13 Fri	10	0			15	9			34	39	11	5	16
14 Sat			169	17			92	105	383	384	97	68	165

Gambar 4.1 Data Awal

Tahapan selanjutnya adalah *preprocessing* adalah pembersihan data (*cleaning*). Pembersihan data bertujuan untuk menangani data yang bernilai kosong (*null*) dan menangani *outlier* yaitu data yang menyimpang secara ekstrim atau nilainya jauh dengan data yang lain. Null adalah representasi dari data kosong. Kehadiran nilai kosong dapat berdampak signifikan untuk perfoma data. Dari total *records* data laporan *player* FO lotuslakes dari masing-masing atribut yaitu 2.249 dan akan dilakukan pembersihan data yang kosong (*missing value*) dengan cara dihapuskan secara manual di Ms. Excel. Penyebab data kosong dikarenakan pada hari senin tidak beroperasi dan tidak ada yang bermain di hari tersebut. Dataset tersebut dapat terlihat pada Gambar 4.2.

Member	Guest	Golf Cart	Rounds
97	85	70	182
45	146	87	200
9	53	34	76
3	21	15	24
6	61	38	79
5	30	18	45
184	220	172	404
154	167	135	321
0	0	0	0
2	74	39	91
8	47	66	31
3	33	42	21
10	24	39	16
186	197	384	165
147	165	315	112
6	101	112	61
6	37	43	26
4	36	44	20



Member	Guest	Golf Cart	Rounds
97	85	70	182
45	146	87	200
9	53	34	76
3	21	15	24
6	61	38	79
5	30	18	45
184	220	172	404
154	167	135	321
2	74	39	91
8	47	66	31
3	33	42	21
10	24	39	16
186	197	384	165
147	165	315	112
6	101	112	61
6	37	43	26
4	36	44	20
3	14	17	10
180	168	348	134

Gambar 4. 2 Pembersihan Data

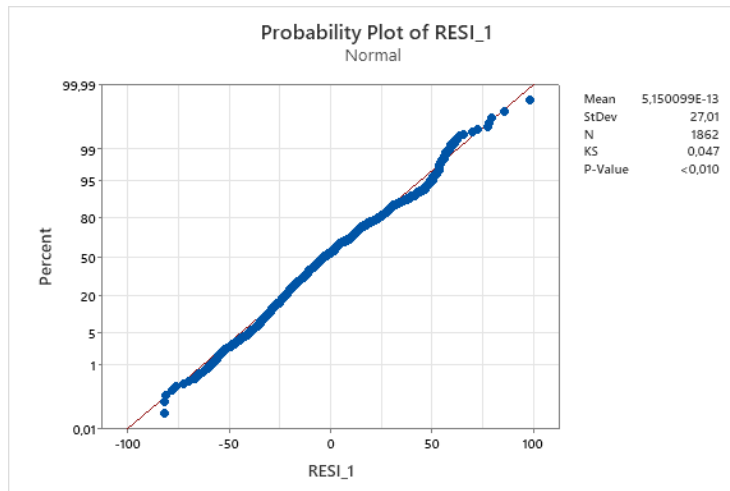
Tahapan selanjutnya adalah Transformasi data. Transformasi data adalah cara yang dilakukan dengan tujuan untuk merubah skala data. Dalam penelitian ini untuk menangani uji asumsi klasik yang hasil uji normalitasnya tidak normal maka

dilakukan transformasi data. Transformasi data ini dilakukan dengan menggunakan bantuan *tools* Minitab 16. Pada tahap ini digunakan transformasi dengan menggunakan logaritma natural semilog, yaitu variabel independen yaitu *member*, *guest* dan *golf cart* ditransformasi dalam bentuk logaritma dan variabel dependen yaitu *round* tetap. Pemilihan atau pentransformasian hanya kepada variabel *member* dan *guest*, dikarenakan 2 variabel ini dianggap memiliki rentang nilai data yang terlalu tinggi. Berikut dapat dilihat pada **Tabel 4.1**

Tabel 4.1 Hasil Transformasi Data

Round	LN_member	Guest	LN_golfcart
182	4,574710979	85	1,44656486
200	3,80666249	146	1,49647258
76	2,197224577	53	1,260266326
24	1,098612289	21	0,996228893
79	1,791759469	61	1,291320319
45	1,609437912	30	1,06138513
404	5,214935758	220	1,638510087
321	5,036952602	167	1,590311112
91	0,693147181	74	1,298435802
66	2,079441542	47	1,233722036
42	1,098612289	33	1,113344054
39	2,302585093	24	1,019781441
384	5,225746674	197	1,63040564
315	4,990432587	165	1,551490713
112	1,791759469	101	1,413635625
43	1,791759469	37	1,181143141
44	1,386294361	36	1,0971887
17	1,098612289	14	0,834032445
348	5,192956851	168	1,588794251
266	4,997212274	117	1,551490713

Setelah dataset ditransformasikan, dataset akan dianalisis dengan menggunakan algoritme regresi linier berganda dengan menggunakan uji asumsi klasik dengan bantuan *tools* Minitab 16. Uji Normalitas dilakukan untuk menguji apakah data tersebut normal atau tidak. Dalam penelitian ini hasil uji normalitas pada **Gambar 4.4**, menunjukkan P-Value < 0,010 yang artinya normal, karena P-Value jika diambil nilai alpha 0,05 maka P-Value < 0,010 < alpha, artinya sebaran data normal. Berikut dapat dilihat pada **Gambar 4.4** dibawah.



Gambar 4.4 Hasil Uji Normalitas

Berdasarkan pengujian dengan teknik visual menunjukkan bahwa titik penyebaran hampir sesuai mendekati garis diagonal dan tidak terlalu melebar menjauh dengan kata lain data menyebar secara linier dan tidak ada data yang *outlier*. Hasil dari uji multikolinieritas, dapat dilihat pada Gambar 4.5 sebagai berikut

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-32,87	2,79	-11,78	0,000	
LN_member	40,197	0,788	51,04	0,000	5,47
Guest	1,4164	0,0227	62,26	0,000	4,06
LN_golfcart	-30,66	3,78	-8,11	0,000	5,37

Gambar 4.5 Hasil Uji Multikolinieritas

Gambar 4.5 menunjukkan nilai VIF untuk variabel member adalah 5,47, guest adalah 4,06, dan golf cart adalah 5,37. Nilai VIF dari ketiga variabel tersebut tidak ada yang lebih besar dari 10, maka bisa dikatakan tidak terjadi multikolinieritas. Dengan demikian, model di atas telah terbebas dari adanya multikolinieritas, karena merujuk pada syarat asumsi klasik regresi linier dengan OLS, maka model regresi linier yang baik adalah yang terbebas dari multikolinieritas.

Untuk menguji autokorelasi dalam uji asumsi klasik memiliki ketentuan sebagai berikut :

Jika nilai $DW < -2$. Terjadi autokorelasi positif

Jika $-2 < DW < 2$. Tidak terjadi autokorelasi.

Jika nilai $DW > 2$. Terjadi autokorelasi negatif.

Dalam penelitian ini, nilai Durbin-watson yang dihasilkan adalah 1,62017 dapat dilihat pada **Gambar 4.6** yang artinya nilai DW berada diantara -2 sampai 2, yang artinya tidak terjadi autokorelasi pada data yang digunakan. Dapat disimpulkan bahwa pengujian ini tidak terjadi autokorelasi.

Durbin-Watson Statistic

Durbin-Watson Statistic = 1,62017

Gambar 4.6 Hasil Durbin-Watson

Pada **Gambar 4.6** dapat diketahui bahwa P-Value dari variabel member sebesar 0,067, variabel guest sebesar 0,877, dan variabel golf cart sebesar 0,154. Jadi dapat diperoleh kesimpulan bahwa semua variabel dalam penelitian ini memiliki P-Value diatas 0,05 maka dapat dinyatakan model regresi ini terbebas dari heteroskedastisitas dan dapat digunakan pengujian selanjutnya.

Pengujian parameter ini bertujuan untuk mengetahui ada atau tidak pengaruh variabel bebas, baik secara serentak maupun secara individu. Uji ini terbagi menjadi dua, yaitu uji f dan uji t.

1. Uji T

Setelah proses uji asumsi klasik dilakukan, data diuji secara parsial dengan menggunakan uji T yang dapat dilihat pada nilai T pada tabel coefficients dalam Minitab lalu dibandingkan dengan nilai pada T-Tabel. Hasil uji T dapat dilihat pada berikut:

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-32,87	2,79	-11,78	0,000	
LN_member	40,197	0,788	51,04	0,000	5,47
Guest	1,4164	0,0227	62,26	0,000	4,06
LN_golfcart	-30,66	3,78	-8,11	0,000	5,37

Gambar 4.7 Hasil Tabel Coefficients

Gambar 4.7 menunjukkan hasil nilai T-Tabel member sebesar 51,04, guest sebesar 62,26, dan golf cart -8,11. Kemudian dapat dicari menggunakan perbandingan antara T-Hitung dengan T-Tabel. T-Tabel dapat dicari menggunakan excel dengan menggunakan rumus $T.INV(probabilitas, degree\ of\ freedom)$ dapat dilihat pada **Gambar 4.8** sebagai berikut :

fx | =T.INV(0,95;1862)

D	E
1,645672	

Gambar 4.8 Hasil T-Tabel

Berdasarkan **Gambar 4.7** dan **Gambar 4.8**. Maka hasil uji t pada penelitian ini dapat dijelaskan sebagai berikut:

1. Member (X_1)

Hipotesis Member adalah :

- H_0 = Tidak terdapat pengaruh member terhadap rounds.
- H_a = Terdapat pengaruh member terhadap rounds.

Perbandingannya adalah T-Hitung pada Minitab 50,04 dan T-Tabel sebesar 1,645672. Jadi $T-Hitung > T-Tabel$ yang dapat disimpulkan bahwa hasil uji T-Hitung $50,04 > 1,645672$. Artinya H_a Terdapat pengaruh member terhadap rounds.

2. Guest (X_2)

Hipotesis Guest adalah :

- H_0 = Tidak terdapat pengaruh guest terhadap rounds.
- H_a = Terdapat pengaruh guest terhadap rounds.

Perbandingannya adalah T-Hitung pada Minitab 62,26 dan T-Tabel sebesar 1,645672. Jadi $T-Hitung > T-Tabel$ yang dapat disimpulkan bahwa hasil uji T-Hitung $62,26 > 1,645672$. Artinya H_a Terdapat pengaruh guest terhadap rounds.

3. Golf Cart (X_3)

Hipotesis Golf Cart adalah :

- Ho = Tidak terdapat pengaruh golf cart terhadap rounds.
- Ha = Terdapat pengaruh golf cart terhadap rounds.

Perbandingannya adalah T-Hitung pada Minitab -8,11 dan T-Tabel sebesar 1,645672. Jadi T-Hitung > T-Tabel yang dapat disimpulkan bahwa hasil uji T-Hitung -8,11 > 1,645672. Artinya Ho Tidak terdapat pengaruh golf cart terhadap rounds.

Uji secara simultan dengan menggunakan uji F yang dapat dilihat pada nilai F pada tabel *Analisis of Variance* dalam Minitab lalu dibandingkan dengan nilai pada F-Tabel. Hasil uji T dapat dilihat pada berikut:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	3	34874774	11624925	15908,12	0,000
hasilLNMember	1	1903387	1903387	2604,69	0,000
Guest	1	2832808	2832808	3876,55	0,000
hasilLNgc2	1	48035	48035	65,73	0,000
Error	1858	1357742	731		
Lack-of-Fit	1467	1357473	925	1348,58	0,000
Pure Error	391	268	1		
Total	1861	36232516			

Gambar 4.9 Hasil Analysis of Variance

Hasil uji F sebesar 15908,12, kemudian dapat dicari menggunakan perbandingan antara F-Hitung dengan F-Tabel. F-Tabel dapat dicari menggunakan excel dengan menggunakan rumus F.INV.RT(*probabilitas, degree of freedom 1, degree of freedom 2*) dapat dilihat pada Gambar 4.10 sebagai berikut :

f_{α} =F.INV.RT(0,5;2;1859)

D	E	F
0,693406		

Gambar 4.10 Hasil F-Tabel

Pada penelitian ini menggunakan analisis regresi linier berganda maka nilai yang digunakan adalah nilai Adjusted R-Square dikarenakan variabel bebas lebih dari satu seperti terlihat pada **Gambar 4.11**

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
27,0325	96,25%	96,25%	96,23%

Gambar 4.11 Hasil Koefisien Determinasi

Gambar 4.11 nilai dari R-sq(adj) yaitu 96,25% menunjukkan bahwa proporsi pengaruh member, guest dan golf cart terhadap rounds sedangkan sisanya 4,75% (100% - 96,25%) dipengaruhi oleh variabel lain yang tidak ada dalam model regresi linier berganda.

Dari hasil regresi dengan menggunakan tools minitab, maka didapatkan model regresi yang dapat dilihat pada **Gambar 4.12**.

Regression Equation

Round = -32,87 + 40,197 LN_member + 1,4164 Guest - 30,66 LN_golfcart

Gambar 4.12 Hasil Regresi Equation

Maka diperoleh hasil persamaan regresi linier berganda yaitu sebagai berikut:

$$\text{Rounds (Y)} = -32,87 + 40,197 \text{ LN_member}(X_1) + 1,4164 \text{ Guest}(X_2) - 30,66 \text{ LN_golfcart}(X_3).$$

Berikut gambaran hitung manual, jika member bernilai 4,574711, guest 85, dan golf cart 1,44656:

$$\begin{aligned} Y &= -32,87 + 40,197(4,574711) + 1,4164(85) - 30,66(1,446565) \\ &= -32,87 + 183,890 + 120,394 - 44,352 \\ &= 227,062 \end{aligned}$$

Jika member bernilai 2,197225, guest 53, dan golf cart 1,260266:

$$\begin{aligned} Y &= -32,87 + 40,197(2,197225) + 1,4164(53) - 30,66(1,260266) \\ &= -32,87 + 88,322 + 75,069 - 38,640 \\ &= 91,881 \end{aligned}$$

Apabila jumlah pemain golf member hari ini 4,574711, guest 85, dan golf cart 1,44656 maka hasil yang diprediksi untuk jumlah roundnya berada di kisaran 227,062. Apabila jumlah pemain golf member hari ini 2,197225, guest 53, dan golf cart 1,260266, maka hasil yang diprediksi untuk jumlah roundnya berada di kisaran 91,881. Angka prediksi tersebut ada dalam output Minitab, seperti pada Tabel 4.2 di bawah ini:

Tabel 4.2 Hasil Prediksi Rounds

Rounds	LN_member	Guest	LN_golfcart	Prediksi	Residual
182	4,574711	85	1,446565	227,0659	-45,0659
200	3,806662	146	1,496473	281,0636	-81,0636
76	2,197225	53	1,260266	91,88321	-15,8832
24	1,098612	21	0,996229	10,49137	13,50863
79	1,791759	61	1,29132	85,9638	-6,9638
45	1,609438	30	1,061385	41,77545	3,224552
404	5,214936	220	1,63851	438,1328	-34,1328
321	5,036953	167	1,590311	357,386	-36,386
91	0,693147	74	1,298436	59,99761	31,00239
66	2,079442	47	1,233722	79,46394	-13,4639
42	1,098612	33	1,113344	23,89782	18,10218
39	2,302585	24	1,019781	62,41523	-23,4152
...

Evaluasi adalah tahapan interpretasi terhadap hasil *final* data mining. Evaluasi dilakukan secara mendalam dengan tujuan agar hasil pada tahap modeling sesuai dengan sasaran yang ingin dicapai. Ada beberapa teknik pengujian yang dilakukan pada penelitian untuk menghitung kesalahan peramalan yaitu *Mean Absolute Deviation* (MAD), *Mean Square Error* (MSE), *Mean Absolute Percentage Error* (MAPE) dan *Root Mean Square Error* (RMSE). Berikut Perhitungan nilai MAD, MSE, MAPE, dan RMSE:

$$\begin{aligned} \text{MAD} &= \frac{\sum |Y^1 - Y|}{n} \\ &= 70,98265 / 1863 \\ &= 0,038101262 \\ \text{MSE} &= \frac{\sum (Y^1 - Y)^2}{n} \\ &= 1351,307149 \end{aligned}$$

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y^1 - Y}{Y^1} \right| \times 100$$

$$= 0,06399372$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum (Y^1 - Y)^2}$$

$$= 36,76012989$$

Tabel 4.3 Hasil Pengujian Evaluasi

MSE	1351,307149
RMSE	36,76012989
MAD	0,038101262
MAPE	0,06399372

Tabel 4.3 menunjukkan bahwa hasil nilai MSE 1351,307149, nilai RMSE sebesar 36,76012989, MAD sebesar 0,038101262, dan MAPE sebesar 0,06399372. Tingkat error dari perhitungan MAD sebesar 0,038101262, yang artinya tingkat error sebesar 3,81% atau < 10% tergolong error kecil atau hasil prediksi sangat akurat. Jika dataset yang digunakan semakin banyak, maka tingkat kesalahan error akan semakin kecil.

PEMBAHASAN

Berdasarkan hasil pengujian yang dilakukan pada data player golf FO di Lotuslakes dengan menggunakan tools Minitab 16 bahwa algoritma Regresi Linier Berganda dapat digunakan untuk melakukan prediksi. Pengujian normalitas menunjukkan P-Value < 0,010 yang artinya tidak normal, karena nilai signifikan harus kurang dari 0,05. Pengujian normalitas data ini menunjukkan P-Value < 0,010 yang artinya normal, karena P-Value jika diambil nilai alpha 0,05 maka P-Value < 0,010 < alpha, artinya sebaran data normal.

Dari hasil uji asumsi klasik yaitu dengan uji multikolinieritas didapatkan bahwa nilai VIF dari 3 variabel X yaitu variabel member adalah 5,47, guest adalah 4,06, dan golf cart adalah 5,37. Dari hasil tersebut tidak ada yang lebih besar dari 10, maka dapat dikatakan tidak terjadi multikolinieritas. Dengan demikian, model yang digunakan pada penelitian telah terbebas dari adanya multikolinieritas.

Selanjutnya hasil uji asumsi klasik dengan uji autokorelasi diperoleh nilai Durbin-watson yang dihasilkan adalah 1,62017 yang artinya nilai DW berada diantara -2 sampai 2, tidak terjadi autokorelasi pada data yang digunakan.

Uji asumsi klasik yaitu uji heteroskedastisitas dengan uji glejser diperoleh hasil bahwa tingkat signifikan variabel member sebesar 0,067, variabel guest sebesar 0,877, dan variabel golf cart sebesar 0,154. Dari hasil uji glejser dapat diperoleh kesimpulan bahwa semua variabel dalam penelitian ini memiliki tingkat signifikan diatas 0,05. Berdasarkan dasar pengambilan keputusan dari uji heteroskedastisitas dengan uji glejser maka dapat dinyatakan model regresi ini terbebas dari heteroskedastisitas dan dapat digunakan pengujian selanjutnya.

Hasil uji kelayakan model dengan uji t secara parsial didapatkan 2 variabel yang mempengaruhi jumlah rounds yaitu variabel member dan guest. Sedangkan variabel golf cart tidak mempengaruhi jumlah rounds. Tetapi uji F secara simultan disimpulkan bahwa hasil uji F-Hitung 15908,12 > 0,693406. Artinya Ha Terdapat pengaruh member, guest dan golf cart secara bersama-sama terhadap rounds.

Selanjutnya hasil koefisien determinasi dari Adjusted R Square yaitu 96,25% menunjukkan bahwa proporsi pengaruh member, guest dan golf cart terhadap rounds sedangkan sisanya 4,75% (100% - 96,25%) dipengaruhi oleh variabel lain yang tidak ada dalam model regresi linier berganda.

Kemudian hasil dari uji koefisien regresi diperoleh hasil persamaan regresi linier berganda yaitu sebagai berikut:

$\text{Rounds (Y)} = -32,87 + 40,197 \text{ LN_member}(X_1) + 1,4164 \text{ Guest}(X_2) - 30,66 \text{ LN_golfcart}(X_3)$. Dari hasil persamaan tersebut maka didapatkan hasil prediksi apabila jumlah pemain golf member hari ini 4,574711, guest 85, dan golf cart 1,44656 maka hasil yang diprediksi untuk jumlah roundsnya berada di kisaran 227,062. Selanjutnya dilakukan tahap evaluasi dengan teknik pengujian yaitu MAD, MSE, MAPE, dan RMSE. Dari keempat hasil evaluasi tersebut didapatkan bahwa nilai MAD mempunyai kemampuan peramalan yang baik dikarenakan nilainya kurang dari 20% yaitu sebesar 0,038101262 atau 3,81%.

SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat disimpulkan sebagai berikut: Teknik regresi linier berganda ini digunakan untuk melakukan pengujian jumlah rounds terhadap member, guest, dan golf cart dengan menggunakan bantuan tools Minitab. Hasil pemodelan menggunakan algoritma ini menghasilkan persamaan regresi yang dapat digunakan sebagai acuan dalam memprediksinya. Teknik regresi linier berganda dapat digunakan untuk mengetahui hasil prediksi jumlah *rounds* sebagai salah satu cara untuk memecahkan permasalahan sulitnya memprediksi jumlah *rounds* karena fluktuasinya yang tinggi.

Hasil evaluasi yang didapatkan dari algoritme regresi linier berganda berupa persamaan regresi yaitu $\text{Rounds (Y)} = -32,87 + 40,197 \text{ LN_member}(X_1) + 1,4164 \text{ Guest}(X_2) - 30,66 \text{ LN_golfcart}(X_3)$. dan hasil evaluasi tingkat error yaitu nilai MAD (*Mean Absolute Deviation*), MSE (*Mean Square error*), MAPE (*Mean Absolute Percent*), dan RMSE (*Root Mean Square Error*), Hasil evaluasi dari MAD yaitu 0,038101262, MSE yaitu 1351,307149, MAPE yaitu 0,06399372, dan RMSE yaitu 36,76012989. Setelah pengujian selesai dapat diketahui bahwa hasil pengaruh member, guest, dan golf cart terhadap rounds signifikan kuat, positif dan searah yaitu sebesar 96,25%, yang artinya setiap kenaikan angka member, guest dan golf cart akan menaikkan prediksi jumlah round dan sebaliknya. Oleh karena itu, management lotuslakes dapat mengambil keputusan yang tepat untuk memaksimalkan keuntungan dan meminimalisir kerugian saat berbisnis, seperti mengadakan event disetiap bulan, atau promo harga dihari-hari tertentu, serta bekerja sama dengan booking golf online, sehingga menarik minat pegolf untuk bermain dan bisa jadi bermain lebih dari satu round.

REFERENCES

- Adams, K. J., Sevene, T., Walsh, J., Climstein, M., & DeBeliso, M. (2020). The golf caddie—the forgotten worker. *J Phys Act Res*, 5(1), 41-44.
- Ali, A. T., Siregar, N. M., & Budiningsih, M. (2022). MODEL PERMAINAN OLAHRAGA GOLF UNTUK ANAK USIA 6-8 TAHUN. *Journal Olahraga Rekat (Rekreasi Masyarakat)*, 12-2, Vol 1 No 1.

- Arhami, M., Kom, M., & Muhammad Nasir, S. T. (2020). *Data Mining-Algoritma dan Implementasi*. Penerbit Andi.
- Asohi Yaya & Andri, A. (2020). Impelementasi Algoritme Regresi Linier Berganda Untuk Prediksi Penjualan. *Jurnal Nasional Ilmu Komputer*, 1(3), 149-158.
- Aqibah, M., Suciptawati, N. L. P., & Sumarjaya, I. W. (2020). Model Dinamis Autoregressive Distributed Lag (Studi Kasus: Pengaruh Kurs Dolar Amerika Dan Inflasi Terhadap Harga Saham Tahun 2014-2018). *E-Jurnal Matematika*, 9(4), 240-250.
- Baihaqi, W. M., Dianingrum, M., & Ramadhan, K. A. N. (2019). Regresi Linier Sederhana Untuk Memprediksi Kunjungan Pasien di Rumah Sakit Berdasarkan Jenis Layanan dan Umur Pasien. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 10(2), 671-680.
- Dewi Rosiana. (2018). *Analisis Algoritma Regresi Linier Berganda Untuk Memprediksi Jumlah Produksi Padi Di Kabupaten Karawang*. Universitas Singaperbangsa Karawang, Karawang.
- Faid, M., Jasri, M., & Rahmawati, T. (2019). Perbandingan Kinerja Tool Data Mining Weka dan Rapidminer Dalam Algoritme Klasifikasi. *Teknika*, 8(1), 11-16.
- Fajrin, A. A., & Maulana, A. (2018). Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritme Fp-Growth Pada Data Transaksi Penjualan Spare Part Motor. *Kumpulan Jurnal Ilmu Komputer (KLIK)*, 5(01), 1-10.
- Firdaus, A., Firdaus, W. I., Studi, P., Informatika, T., Digital, M., & Sriwijaya, P. N. (2021). Text mining. *Jurnal Penelitian Ilmu dan Teknologi Komputer*, 13(1), 66–78.
- Gaol, I. L. L., Sinurat, S., & Siagian, E. R. (2019). Implementasi Data Mining Dengan Metode Regresi Linear Berganda Untuk Memprediksi Data Persediaan Buku Pada Pt. Yudhistira Ghalia Indonesia Area Sumatera Utara. *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, 3(1).
- Harlan, J. (2018). *Analisis Regresi Linier*. Jakarta: Gunadarma.
- Irianto, T. (2020). *Olahraga*. Radar Banjarmasin. Retrieved from <http://eprints.ulm.ac.id/id/eprint/5723>
- Katik, I., & Putro, H. P. H. (2019). Studi Karakteristik, Motivasi dan Penilaian Golfer pada Golf Course sebagai Destinasi Wisata di Bandung dan Sekitarnya. *Jurnal Perencanaan Kepariwisata*.
- Malik, S. M., Taurusyanti, D., & Yusnita, N. (2021). Hubungan disiplin kerja dengan kinerja karyawan di perusahaan cnn indonesia tv bagian marketing pr head. *Jurnal Online Mahasiswa (JOM) Bidang Manajemen*, 6(3). 1-10.
- Nangi, J., Indrianti, S. H., & Pramono, B. (2018). Peramalan persediaan obat menggunakan metode Triple Exponential Smoothing (Tes) (studi kasus: Instalasi Farmasi rsud kab. Muna). Vol.4 No.1, Hal 135-142.
- Padilah, T. N., & Adam, R. I. (2019). Analisis regresi linier berganda dalam estimasi produktivitas tanaman padi di Kabupaten Karawang. *FIBONACCI: Jurnal Pendidikan Matematika Dan Matematika*, 5(2), 117-128.
- Purwanto, Purwanto. (2019). Analisis korelasi dan regresi linier dengan spss 21 (panduan praktis untuk penelitian ekonomi syariah). Purworejo : StaiaPress.

- Puteri, K., & Silvanie, A. (2020). Machine Learning Untuk Model Prediksi Harga Sembako Dengan Metode Regresi Linear Berganda. *Jurnal Nasional Informatika (JUNIF)*, 1(2), 82-94.
- Putra, R. A., & Martha, L. (2022). Pengaruh Keselamatan Kerja, Kesehatan Kerja dan Lingkungan Kerja Terhadap Produktivitas Kerja Karyawan Pada PT. Bima Arjuna Prakasa Padang. *Jurnal Valuasi: Jurnal Ilmiah Ilmu Manajemen dan Kewirausahaan*, 2(2), 1061-1084.
- Rachmawati, R. I., Kriswantoro, K., & Nasution, M. (2019). Manajemen Klub Golf Semarang Olympic Golf Club (SOGC) Kota Semarang Tahun 2017. *Journal of Sport Coaching and Physical Education*, 4(1), 21-27.
- Rizky, F., Syahra, Y., & Mariami, I. (2019). Implementasi Data Mining Untuk Memprediksi Target Pemakaian Stok Barang Menggunakan Metode Regresi Linier Berganda. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika dan Komputer)*, 18(2), 167-175.
- Sari, Y. R., Ananda, L. R., & Rani, M. (2021). Perbandingan Analisis Logika Fuzzy dan Regresi Linier Berganda dalam Menentukan Produksi Beras Nasional. *Jurnal KomtekInfo*, 239-248.
- Sella Bela Tiara. (2020). *Prediksi Ketersediaan Stok Obat Dengan Menggunakan Algoritme Regresi Berganda (Studi Kasus: Puskesmas Purwasari)*. Universitas Singaperbangsa Karawang, Karawang.
- Siregar, A. Z. (2021). Implementasi Metode Regresi Linier Berganda Dalam Estimasi Tingkat Pendaftaran Mahasiswa Baru. *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, 2(3), 133-137.
- Sulistiyono, S., & Sulistiyowati, W. (2018). Peramalan produksi dengan metode regresi linier berganda. *PROZIMA (Productivity, Optimization and Manufacturing System Engineering)*, 1(2), 82-89.
- Wulandari, N. A., Pratiwi, H., & Handayani, S. S. (2023, Januari). Perbandingan Metode K-means and Agglomerative Nesting untuk Clustering Data Digital Marketing di Twitter. In *STAINS (SEMINAR NASIONAL TEKNOLOGI & SAINS)* (Vol. 2, No. 1, pp. 189-194).
- _____. *Undang-undang No. 3 Tahun 2005 Pasal 1 Ayat 4 tentang Sistem Keolahragaan Nasional*
- _____. *Undang-undang No. 3 Tahun 2005 Pasal 17 tentang Sistem Keolahragaan Nasional*