

Prediksi Pertumbuhan Jumlah Unit Usaha Mikro Kecil dan Menengah (UMKM) Menggunakan Metode Regresi Linier Berganda

¹Firmansyah, ² Agus Yulianto
Universitas Bina Sarana Informatika
Jakarta, Indonesia
Universitas Nusa Mandiri
Jakarta, Indonesia

firmansyah.fmh@bsi.ac.id, agus.asy@nusamandiri.ac.id

*Corresponding Author

Diajukan : 08/04/2022
Diterima : 12/04/2022
Dipublikasi : 13/04/2022

ABSTRAK

Dengan pesatnya pertumbuhan UMKM (Usaha Mikro Kecil dan Menengah) di Indonesia namun belum adanya penelitian yang memprediksi pertumbuhan UMKM. Dengan memprediksi pertumbuhan UMKM maka pemerintah juga dapat menentukan kebijakan-kebijakan seperti stimulus, pemberdayaan ekonomi dan juga dapat memproyeksikan pertumbuhan ekonomi. Prediksi menggunakan model regresi linier berganda dengan menggunakan 2 variabel dependent dan 1 variabel independen. Data diambil dari jumlah pertumbuhan UMKM dan suku bunga bank. Hasil pemodelan membuktikan bahwa model regresi linier berganda dapat memprediksi pertumbuhan UMKM.

Kata Kunci: UMKM, Pertumbuhan Ekonomi, Regresi Linear, Prediksi Pertumbuhan UMKM

I. PENDAHULUAN

Pertumbuhan Usaha Mikro Kecil dan Menengah (UMKM) diiringi juga dengan mulai membaiknya pertumbuhan ekonomi Indonesia. Semakin tingginya pertumbuhan UMKM tentunya akan menyumbang pendapatan domestik bruto (PDB) Indonesia. Sampai dengan tahun 2019, berdasarkan data Kementerian Koperasi dan UKM (KemenkopUKM) bahwa UMKM berkontribusi 60,51% terhadap PDB (KemenkopUKM, 2022).

Dengan pesatnya pertumbuhan UMKM di Indonesia namun belum adanya penelitian yang memprediksi pertumbuhan UMKM. Dengan memprediksi pertumbuhan UMKM maka pemerintah juga dapat menentukan kebijakan-kebijakan seperti stimulus, pemberdayaan ekonomi dan juga dapat memproyeksikan pertumbuhan ekonomi.

Untuk memprediksi pertumbuhan UMKM maka digunakan metode regresi linier berganda berdasarkan indikator jumlah unit usaha. Dengan regresi linier berganda, maka algoritma dapat memprediksi jumlah unit usaha pada periode tahun berikutnya

II. STUDI LITERATUR

1. Penelitian Terdahulu

Penelitian yang terkait dengan metode regresi linier sudah dilakukan oleh peneliti-peneliti sebelumnya seperti memprediksi penjualan sepatu yang paling laku dengan menggunakan regresi linier (Morinatinambunan & Hasugian, 2020) dan memprediksi dampak dari iklan di Youtube dalam mempengaruhi penjualan (Zhou et al., 2021). Dalam permasalahan bisnis, regresi linier berganda digunakan untuk memprediksi hasil bisnis (Rusov et al., 2017), membuat forecast untuk produk yang cepat terjual menggunakan regresi linier berganda (Farizal et al., 2020). Penelitian yang terkait edukasi yaitu estimasi kelulusan siswa (Kurniatullah & Pramudi, 2017),

memprediksi performa siswa menggunakan regresi linier berganda (Tamrakar, 2021) dan memprediksi tingkat kepuasan siswa dalam proses belajar jarak jauh selama pandemic COVID-19 (Ho et al., 2021). Dalam penelitian organisasi, digunakan untuk memprediksi kepuasan karyawan atas pekerjaan (Ramesh Tharu, 2019). Regresi linier juga digunakan untuk memproyeksikan permintaan listrik di Colombia (Garcia-Guilianny et al., 2020), mengukur kinerja energi di lingkungan tempat tinggal dengan menggabungkan dengan metode Bayesian (Permai & Tanty, 2018) dan memprediksi konsumsi energi pada fasilitas kolam renang (Smedegård et al., 2021). Untuk penelitian yang terkait dengan lingkungan, regresi linier berganda digunakan untuk mengestimasi indeks kualitas air yang digabungkan dengan AI (Gaya et al., 2020). Penelitian untuk memprediksi kehilangan pelanggan atau customer churn menggunakan algoritma Naïve Bayes, dari hasil pemodelan yang dilakukan Naïve Bayes terbukti dapat memprediksi customer churn pada bisnis retail dengan variabel jenis kelamin, grade point, kepemilikan kartu kredit, rentang usia dan nilai rata-rata transaksi (Firmansyah & Yulianto, 2021).

2. UMKM

Usaha Mikro Kecil Menengah (UMKM) merupakan unit usaha produktif yang berdiri sendiri, yang dilakukan oleh orang perorangan atau badan usaha disemua sektor ekonomi (Tambunan, 2017).

3. Regresi Linier

Metode regresi linier adalah sebuah metode statistik yang melakukan prediksi menggunakan pengembangan hubungan matematis antar variabel, yaitu variabel dependen (Y) dengan variabel independen (X). Prediksi terhadap nilai variabel dependen dapat dilakukan apabila variabel independennya diketahui. Rumus untuk regresi linier berganda menggunakan persamaan di bawah ini :

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$y = a + b.x$$

Keterangan :

a, b = konstanta

x = variable dependent

y = variable independent

n = jumlah data

III. METODE

A. CRISP-DM (Cross-Industry Standard Process for Data Mining)

Framwork yang digunakan dalam implementasi data mining menggunakan CRISP-DM (Cross-Industry Standard Process for Data Mining) merupakan standar metode implementasi data mining untuk industry (Larose, 2006), adapun prosesnya yaitu :

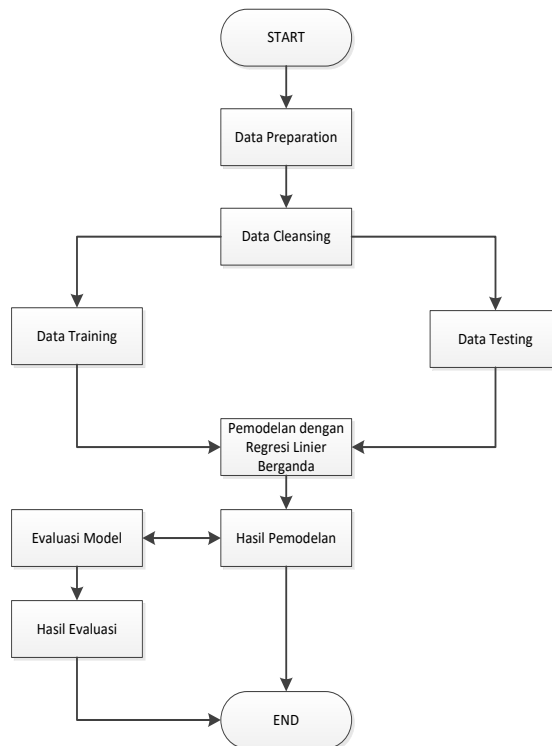
1. Business Understanding Phase
2. Data Understanding Phase
3. Data Preparation Phase
4. Modelling Phase
5. Evaluation Phase
6. Deployment Phase



Gambar 1 CRISP-DM Process

B. Flowchart Pemodelan Regresi Linier Berganda

Flowchart digunakan untuk menggambarkan alur proses mulai dari data preparation hingga pemodelan. Flowchart untuk menggambarkan proses pemodelan menggunakan metode regresi linier berganda ini seperti di bawah ini :



Gambar 2 Proses Pemodelan

IV. HASIL DAN PEMBAHASAN

Data merupakan data sekunder yang diambil dari Badan Pusat Statistik dan KemenkopUKM periode 2010-2019. Data yang diambil yaitu rata-rata suku bunga bank persero dan inflasi periode tahun 2010-2019 bersumber dari Badan Pusat Statistik (BPS), sedangkan angka pertumbuhan UMKM diambil dari Kementerian Koperasi dan UMKM mulai tahun 2010-2019.

A. Data Training dan Data Testing

Data training berjumlah 10 baris dan 3 atribut dengan klasifikasi seperti di bawah ini :

Tabel 1 Klasifikasi Variabel

Variable name	Tipe data
Rata-rata suku bunga bank persero (X1)	Numerik (persentase)
Inflasi (X2)	Numerik (persentase)
Pertumbuhan UMKM (Y)	Numerik

Dari hasil pemodelan data training, maka nanti akan diperoleh persamaan regresi linier yang kemudian akan dapat digunakan untuk memasukkan variabel testing.

B. Regresi Linier Berganda

Data rata-rata suku bunga, inflasi dan pertumbuhan UMKM digabung menjadi satu tabel, khusus data pertumbuhan UMKM nilainya dibagi menjadi per 10.000 agar nilainya menjadi lebih kecil.

Data dibuatkan dalam bentuk spreadsheet dimana ada 2 variabel dependent (X1 dan X2) dan variabel independent (Y) dengan periode tahun 2010 sampai dengan 2019. Masing-masing kolom dihitung jumlah dan rata-rata seperti tabel di bawah ini :

Tabel 2 Perhitungan variabel x dan y

TAHUN	RATA2 SUKU BUNGA BANK PERSERO (X1)	INFLASI (X2)	PERTUMBUHAN UMKM (Y)
2010	13.06	6.96	5,276
2011	12.37	3.79	5,411
2012	11.99	4.3	5,521
2013	11.79	8.38	5,653
2014	12.32	8.36	5,790
2015	12.52	3.35	5,926
2016	11.55	3.02	6,166
2017	10.79	3.61	6,292
2018	10.32	3.13	6,419
2019	10.01	2.72	6,547
Jumlah	116.72	47.62	59,001.78
Rata-rata	11.67	4.76	5,900.18

Langkah berikutnya adalah menghitung masing-masing variabel seperti di bawah ini :

Tabel 3 Perhitungan variabel bagian 1

THN	$y_i = Y_i - \bar{Y}$	$x_{1i} = X_{1i} - \bar{X}_1$	$x_{2i} = X_{2i} - \bar{X}_2$
2010	-623.70	1.39	2.20
2011	-488.70	0.70	-0.97
2012	-379.53	0.32	-0.46
2013	-246.72	0.12	3.62
2014	-110.61	0.65	3.60
2015	26.10	0.85	-1.41
2016	265.48	-0.12	-1.74
2017	392.08	-0.88	-1.15
2018	519.23	-1.35	-1.63
2019	646.37	-1.66	-2.04
SUM	0	0	0

Tabel 4 Perhitungan variabel bagian 2

THN	y_i^2	X_{1i}^2	X_{2i}^2
2010	389,005.66	1.93	4.83
2011	238,823.86	0.49	0.94
2012	144,045.89	0.10	0.21
2013	60,870.26	0.01	13.09
2014	12,233.70	0.42	12.95
2015	681.16	0.72	1.99
2016	70,477.78	0.01	3.03
2017	153,729.49	0.78	1.33
2018	269,597.22	1.83	2.66
2019	417,796.14	2.76	4.17
SUM	175,7261.15	9.05	45.21

Tabel 5 Perhitungan variabel bagian 3

THN	$y_i X_{1i}$	$y_i X_{2i}$	$X_{1i} X_{2i}$
2010	-865.70	-1,370.90	3.05
2011	-341.11	475.01	-0.68
2012	-120.69	175.34	-0.15
2013	-29.11	-892.63	0.43

2014	-71.67	-397.96	2.33
2015	22.13	-36.85	-1.20
2016	-32.39	-462.46	0.21
2017	-345.82	-451.68	1.02
2018	-702.00	-847.38	2.21
2019	-1,074.27	-1,319.89	3.39
SUM	-3,560.63	-5,129.39	10.62

Setelah menghitung variabel, maka selanjutnya adalah menghitung nilai koefisien b1, b2 dan b0 dimana b1 dan b2 merpresentasikan X1 dan X2 sedangkan b0 adalah intercept, untuk menghitung b1 menggunakan persamaan berikut :

$$\hat{b}_1 = \frac{(\sum x_1 y)(\sum x_2^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$\hat{b}_1 = \frac{(-3560.63 * 45.21) - (-5129.39 * 10.62)}{(9.05 * 45.21) - 112.69}$$

$$\hat{b}_1 = -359.28$$

Untuk menghitung b2 menggunakan persamaan berikut :

$$\hat{b}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$\hat{b}_2 = \frac{(-5129.39 * 9.05) - (-3560.63 * 10.62)}{(9.05 * 45.21) - 112.69}$$

$$\hat{b}_2 = -29.10$$

untuk menghitung nilai b0 menggunakan persamaan berikut :

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2$$

$$\hat{b}_0 = -5900 - (-359.28 * 11.67) - (-29.10 * 4.76)$$

$$\hat{b}_0 = 10232.26$$

Kemudian menghitung R square dengan persamaan berikut :

$$R^2 = \frac{\hat{b}_1(\sum x_2 y) + \hat{b}_2(\sum x_1 y)}{\sum y^2}$$

$$R^2 = \frac{(-359.28) * (-3560.63) + (-29.10) * (-5129.39)}{1757261.15}$$

Menentukan nilai Y yang nantinya akan menjadi persamaan untuk prediksi.

$$Y = \hat{b}_0 + \hat{b}_1X_1 + \hat{b}_2X_2$$

$$Y = 10232.26 + (-359.28 * X1) + (-29.10 + X2)$$

Persamaan Y di atas merupakan persamaan yang bisa dijadikan sebagai prediksi, cukup memasukkan nilai X1 dan X2. Misal nilai X1 yaitu 11 dan nilai X2 yaitu 4 sehingga jika dimasukkan ke dalam persamaan Y yaitu :

$$Y = 10232.26 + (-359.28 * X1) + (-29.10 + X2)$$

$$Y = 10232.26 + (-359.28 * 11) + (-29.10 + 4)$$

$$Y = 6163.79$$

Di bawah ini merupakan statistik regresi dan koefisien regresi yang direpresentasikan dalam tabel :

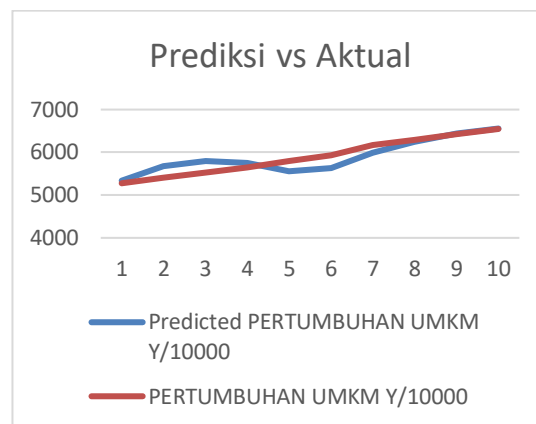
Tabel 6 Statistik Regresi

<i>Regression Statistics</i>	
Multiple R	0.901619105
R Square	0.812917011
Adjusted R Square	0.759464728
Standard Error	216.7136587
Observations	10

Tabel 7 Koefisien Regresi

<i>Coefficients</i>	
Intercept	10232.26384
RATA2 SUKU BUNGA BANK PERSERO (X1)	-359.2816173
INFLASI (X2)	-29.09504837

Untuk mempermudah melihat hasil prediksi dan aktual pada prediksi pertumbuhan UMKM, maka dibuat grafik prediksi versus aktual seperti di bawah ini :



Gambar 3 Grafik prediksi vs aktual

Untuk pengujian akurasi menggunakan model RMSE (*Root Mean Square Error*), MSE (*Mean Square Error*) dan MAPE (*Mean Percentage Absolute Error*). Hasil perhitungannya dalam tabel di bawah ini :

Tabel 8 Pengujian Model

RMSE	181.3156554
MSE	32875.3669
MAPE	5.772435991

V. KESIMPULAN

Model regresi linier berganda terbukti dapat memprediksi pertumbuhan UMKM dengan variabel dependent rata-rata suku bunga dan inflasi. Dari nilai koefisien regresi dapat disimpulkan jika terjadi kenaikan suku bunga dan inflasi maka akan mengurangi pertumbuhan jumlah UMKM. Hasil pengujian model menggunakan MAPE masih dalam kategori sangat baik karena masih dalam nilai dibawah 10% yaitu 5,77%. Model regresi linier berganda cukup akurat dalam memprediksi pertumbuhan UMKM, model ini juga dapat digunakan untuk memprediksi permasalahan lainnya seperti prediksi penjualan, prediksi strategi promosi dan lain-lain. Variabel dependent dalam regresi linier berganda dapat ditambahkan sesuai kebutuhan.

VI. REFERENSI

- Farizal, Qaradhawi, Y., Cornelis, C. I., & Dachyar, M. (2020). Fast Moving Product Demand Forecasting Model With Multi Linear Regression. *AIP Conference Proceedings*, 2227(May). <https://doi.org/10.1063/5.0001031>
- Firmansyah, & Yulianto, A. (2021). Market Basket Analysis for Books Sales Promotion using FP Growth Algorithm, Case Study : Gramedia Matraman Jakarta. *Journal of Informatics Anf Telecommunication Engineering*, 4(January), 383–392.
- Garcia-Guiliany, J., De-La-hoz-franco, E., Rodríguez-Toscano, A. D., De-La-hoz-hernández, J. D., & Hernandez-Palma, H. G. (2020). Multiple Linear Regression Model Applied to The Projection of Electricity Demand in Colombia. *International Journal of Energy Economics and Policy*, 10(1), 419–422. <https://doi.org/10.32479/ijeeep.7813>
- Gaya, M. S., Abba, S. I., Abdu, A. M., Tukur, A. I., Saleh, M. A., Esmaili, P., & Wahab, N. A.

- (2020). Estimation of Water Quality Index Using Artificial Intelligence Approaches and Multi-linear Regression. *IAES International Journal of Artificial Intelligence*, 9(1), 126–134. <https://doi.org/10.11591/ijai.v9.i1.pp126-134>
- Ho, I. M. K., Cheong, K. Y., & Weldon, A. (2021). Predicting Student Satisfaction of Emergency Remote Learning in Higher Education During COVID-19 Using Machine Learning Techniques. *PLoS ONE*, 16(4 April), 1–27. <https://doi.org/10.1371/journal.pone.0249423>
- KemenkopUKM. (2022). *Kementrian Koperasi dan UMKM*. <https://kemenkopukm.go.id/data-umkm>
- Kurniatullah, B. D. F., & Pramudi, Y. T. C. (2017). Estimation of Students' Graduation Using Multiple Linear Regression Method. *Journal of Applied Intelligent System*, 2(1), 29–36. <https://doi.org/10.33633/jais.v2i1.1415>
- Larose, D. T. (2006). *Data Mining Methods and Models*. Johns Wiley & Sons.
- Morinatinambunan, A., & Hasugian, P. M. (2020). *Application of Data Mining to Predict the Sales of the Best-Selling Shoe Products using the Simple Linear Regression Method*. 14(2), 113–120.
- Permai, S. D., & Tanty, H. (2018). Linear Regression Model Using Bayesian Approach for Energy Performance of Residential Building. *Procedia Computer Science*, 135, 671–677. <https://doi.org/10.1016/j.procs.2018.08.219>
- Ramesh Tharu. (2019). Multiple Regression Model Fitted for Job Satisfaction of Employees Working in Saving and Cooperative Organization Multiple Regression Model Fitted for Job Satisfaction of Employees Working in Saving and Cooperative Organization View project Multiple regre. *International Journal of Statistics and Applied Mathematics*, 4(4), 43–49. www.deoc.gov.np/downloadfile/Coop
- Rusov, J., Misita, M., Milanovic, D. D., & Milanovic, D. L. (2017). Applying Regression Models to Predict Business Results. *FME Transactions*, 45(1), 198–202. <https://doi.org/10.5937/fmet1701198R>
- Smedegård, O. Ø., Jonsson, T., Aas, B., Stene, J., Georges, L., & Carlucci, S. (2021). The implementation of multiple linear regression for swimming pool facilities: Case study at Jøa, Norway. *Energies*, 14(16), 1–23. <https://doi.org/10.3390/en14164825>
- Tambunan, T. (2017). *Usaha Mikro Kecil dan Menengah di Indonesia*.
- Tamrakar, A. (2021). Student Performance Prediction by means of Multiple Regression. *Turkish Journal of Computer and Mathematics ...*, 12(12), 4126–4132. <https://www.turcomat.org/index.php/turkbilmart/article/view/8249>
- Zhou, Y., Ahmad, Z., Alsuhabi, H., Yusuf, M., Alkhairy, I., & Sharawy, A. M. (2021). Impact of YouTube Advertising on Sales with Regression Analysis and Statistical Modeling: Usefulness of Online Media in Business. *Computational Intelligence and Neuroscience*, 2021. <https://doi.org/10.1155/2021/9863155>