

Prediksi Penjualan Menggunakan Algoritma Regresi Linear pada Koperasi Karyawan Usaha Bersama

Muhamad Galih¹, Prima Dina Atika^{1,*}, Mukhlis¹

¹ Informatika; Universitas Bhayangkara Jakarta Raya; Jl. Raya Perjuangan No.81 Margamulya, Kota Bekasi, Jawa Barat; e-mail: muhamad.galih18@mhs.ubharajaya.ac.id, prima.dina@dns.ubharajaya.ac.id, mukhlis@dns.ubharajaya.ac.id

* Korespondensi: e-mail: prima.dina@dns.ubharajaya.ac.id

Diterima: 12 Jul 2022; Review: 31 Des 2022; Disetujui: 31 Des 2022; Diterbitkan: 31 Des 2022

Abstract

Koperasi Karyawan Usaha Bersama is one of the cooperative that has minimarket as one of their business strategy that sales a lot of goods people need. The amount of item sold every day is sometimes uncertain. There is a day when a lot of things are sold on that day, and there is also a day when there are only a few stuff are sold. The purpose of this study is to do prediction the sales of goods sold every day. With this prediction result, the minimarket can determine and choose what strategies are the best to maximize their profit. We use Linear Regression to do this study with CRISP-DM as the model process. The result of the analysis that uses data from March to April 2022 shows that MAPE has the better result with 10.7% compared to RMSE with 42.091.

Keywords: linear regression, MAPE, predicting sales, RMSE

Abstrak

Koperasi Karyawan Usaha Bersama adalah salah satu koperasi yang memiliki usaha minimarket yang menjual berbagai kebutuhan pokok yang dibutuhkan masyarakat. Jumlah penjualan barang setiap harinya terkadang tidak menentu, ada hari dimana penjualan sangat tinggi dan penjualan sangat rendah keesokan harinya. Tujuan dari penelitian ini adalah untuk melakukan prediksi terhadap penjualan barang yang terjual setiap harinya. Dengan hasil prediksi ini, pihak koperasi dapat membuat dan menyusun strategi penjualan agar keuntungan yang didapat menjadi maksimal. Penelitian ini dilakukan dengan menggunakan metode Regresi Linear dan juga menggunakan model proses CRISP-DM sebagai alur proses *data mining*. Hasil yang didapatkan dari penelitian ini melibatkan data penjualan dari bulan Maret sampai April 2022 mendapatkan kesalahan sebesar 10,7% pada MAPE dan 42,091 pada RMSE.

Kata kunci: prediksi penjualan, MAPE, regresi linear, RMSE

1. Pendahuluan

Penjualan merupakan suatu fungsi dari pemasaran yang sangat penting dan menentukan bagi perusahaan untuk mencapai tujuan dari perusahaan tersebut yaitu memperoleh laba untuk kelangsungan hidup dari perusahaan tersebut. Dunia bisnis yang terus berubah dengan cepat, serta makin banyaknya persaingan, menuntut perusahaan untuk dapat menganalisis lingkungan bisnis dan memprediksi berbagai kemungkinan yang akan terjadi di masa depan. Peramalan atau forecasting merupakan salah satu usaha perusahaan sebagai dasar pengambilan suatu keputusan yang strategis dalam kelangsungan usaha. Perusahaan pemasaran yang baik

menginginkan informasi untuk membantu mereka mengevaluasi kinerja masa lalu dan merencanakan masa depan, salah satu contoh instansi yang membutuhkan strategi pemasaran yang baik adalah koperasi.

Koperasi adalah sebuah badan usaha yang organisasinya didirikan khusus bertujuan untuk kesejahteraan pada anggotanya (Peraturan Pemerintah Republik Indonesia, 2012). Demikian dengan Koperasi Karyawan Usaha Bersama yang didirikan oleh karyawan sebuah perusahaan swasta dengan tujuan utama untuk meningkatkan kesejahteraan anggotanya. Salah satu usaha yang dimiliki saat ini adalah minimarket yang menjual kebutuhan pokok seperti sembako dan lain sebagainya. Berikut adalah tren penjualan barang yang terjual selama 50 di bulan Maret dan April tahun 2022. Prediksi penjualan barang yang terjual di Koperasi Karyawan Usaha Bersama yang akurat dapat membantu pihak koperasi dalam menentukan strategi penjualan yang baik untuk kedepannya.

Regresi linear melakukan peramalan atau prediksi tentang karakteristik kualitas maupun kuantitas. Hal ini dikarenakan dengan memperkirakan berbagai kombinasi produk, perusahaan dapat memaksimalkan keuntungan serta memperkirakan jumlah produksi yang tepat (Ayuni & Fitriana, 2019).

Penelitian terkait dengan prediksi, diantaranya memprediksi turun atau naiknya tingkat produksi kopi yang melibatkan 5 periode produksi nilai tertinggi tahun 2015 sebesar 1.537,38 ton dan nilai terendah tahun 2011 sebesar 1.109 ton, pengujian MSE nilai 43,112% dan MAPE 20,001% sehingga pengujian MAPE lebih baik dalam menghitung akurasi prediksi produksi kopi (Katemba & Djoh, 2017). Penurunan penjualan supermarket semenjak pandemi, hasil analisa menunjukkan bahwa metode *Moving Average* mempunyai kinerja yang paling baik dengan nilai MSE dan RMSE dari data elektronik adalah 57,603 dan 7,59, data kesehatan adalah 50.489, dan 7.106. Sedangkan metode Regresi Linear memiliki hasil MSE 114,79, dan 10,71, sedangkan RMSE 59.965, dan 7.774 (Nafi'iyah & Rakhmawati, 2021)

2. Metode Penelitian

Proses menemukan korelasi, pola, dan tren baru dari data yang disimpan dalam repositori, menggunakan teknologi pengenalan pola serta statistik dan teknik matematika (Larose, 2005). Penjualan dalam sebuah perusahaan menjadi indikator, bila penjualan besar maka laba dihasilkan pun besar (Pakaja & Naba, 2018).

2.1. Prediksi

Prediksi atau peramalan dapat diasumsikan sebagai kegiatan untuk memperkirakan apa yang akan terjadi di masa yang akan datang (Prasetyowati, 2018). Kegiatan ini dilakukan dengan memperhatikan data atau informasi masa lalu atau saat ini baik secara matematik maupun statistik (Hakimah et al., 2015). Prediksi bertujuan untuk mengetahui, melihat, dan memperkirakan prospek ekonomi atau kegiatan usaha (Ong, 2009).

Sebuah prediksi dapat bersifat kualitatif (tidak berbentuk angka) dan kuantitatif (berbentuk angka) (Syafuruddin et al., 2014). Predikisi bersifat kualitatif cenderung lebih sulit untuk

diprediksi karena variabelnya yang mempunyai sifat sangat relatif (Gustriansyah, 2017). Sedangkan prediksi yang bersifat kuantitatif hasil prediksi yang dibuat sangat bergantung pada metode yang digunakan (Hakimah et al., 2015).

Dalam perancangan suatu metode prediksi atau peramalan, terdapat tiga tahapan yang harus dilalui, yaitu (Indarwati et al., 2019):

- a. Melakukan analisis terhadap data di masa lampau.
- b. Memilih metode yang akan digunakan. Terdapat bermacam-macam metode yang dapat digunakan untuk melakukan peramalan. Pemilihan metode dapat mempengaruhi hasil prediksi atau peramalan. Hasil ramalan diukur menggunakan *error* atau kesalahan terkecil. Oleh karena itu, tidak ada metode peramalan yang pasti baik untuk semua jenis data.
- c. Proses transformasi dari data masa lampau dengan menggunakan metode yang dipilih. Apabila diperlukan maka diadakan perubahan sesuai kebutuhannya.

2.2. Regresi Linear

Regresi Linear adalah metode untuk menyelidiki hubungan antara satu variabel terikat dengan satu variabel bebas. Regresi digunakan untuk mengukur hubungan dua variabel atau lebih yang dinyatakan dengan bentuk hubungan dan fungsi. Regresi Linear ialah bentuk hubungan dimana variabel bebas X maupun Variabel tergantung Y sebagai faktor yang berpangkat satu. Regresi Linear sering digunakan adalah regresi linear sederhana dengan bentuk fungsi pada rumus 1.

$$Y = a + bX \tag{1}$$

Dimana :

Y = Variabel dependen

X = Variabel independen

a = Konstanta / *Intercept*

b = Koefisien regresi / *Slope*

Dengan :

$$a = \frac{\sum Y \sum X^2 - \sum X \sum XY}{n \sum X^2 - (\sum X)^2} \tag{2}$$

$$b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} \tag{3}$$

2.3. Pengujian Akurasi

Kesalahan dalam perumusan prediksi disebabkan ketidakmampuan model peramalan mengenali unsur lain dalam deret data. Diantara beberapa cara untuk menghitung besarnya kesalahan, beberapa diantaranya adalah *root mean squared error* (RMSE), dan *mean absolute percentage error* (MAPE). RMSE metode mengevaluasi teknik peramalan nilai rata-rata dari jumlah kuadrat kesalahan, juga menyatakan besarnya kesalahan suatu model prakiraan. Nilai RMSE rendah suatu model prakiraan mendekati variasi nilai observasinya.

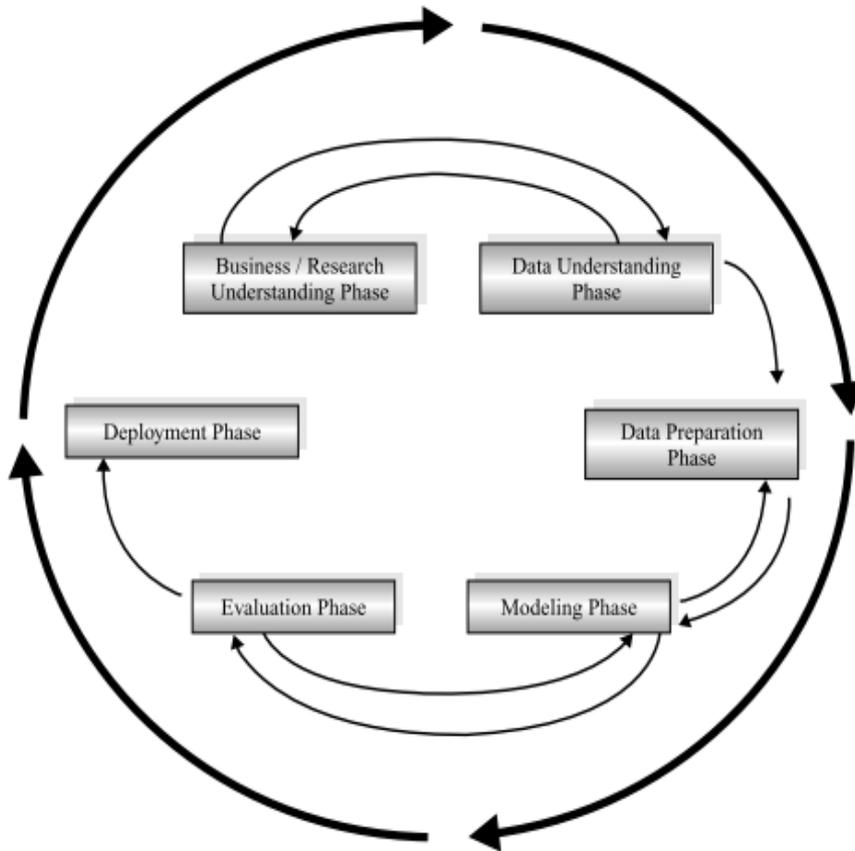
$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}} \quad (4)$$

Sedangkan MAPE adalah merupakan ukuran kesalahan relatif. MAPE menyatakan persentase kesalahan hasil peramalan berupa informasi persentase kesalahan terlalu tinggi atau rendah. Secara matematis, MAPE dinyatakan pada rumus 5.

$$MAPE = \frac{\sum_{t=1}^n \left| \left(\frac{A_t - F_t}{A_t} \right) 100 \right|}{n} \quad (5)$$

2.4. Model Proses *Data mining* Menggunakan CRISP-DM

Cross Industry Standard Process for Data mining atau disingkat CRISP-DM dikembangkan pada tahun 1996 oleh beberapa analis yang mewakili DaimlerChrysler, SPSS, and NCR. CRISP menyediakan sebuah proses standar yang tidak memiliki dan tersedia secara bebas untuk menyesuaikan *data mining* ke dalam strategi pemecahan masalah umum dari unit bisnis atau penelitian (Larose, 2005).



Sumber: Larose (2005)

Gambar 1. CRISP-DM

a. *Bussiness Understanding*

Tahap pertama dalam CRISP-DM dan termasuk bagian yang cukup vital. Pada tahap ini membutuhkan pengetahuan dari objek bisnis, bagaimana membangun atau mendapatkan data, dan bagaimana untuk mencocokkan tujuan pemodelan untuk tujuan bisnis sehingga model terbaik dapat dibangun.

b. *Data Understanding*

Kegiatan untuk mengenal data atau untuk mendeteksi himpunan bagian yang menarik untuk membentuk hipotesis untuk informasi tersembunyi.

c. *Data Preparation*

Kegiatan untuk membangun kumpulan data akhir termasuk tabel, pemilihan atribut serta pembersihan data untuk pemodelan.

d. *Modelling*

Pemilihan teknik pemodelan dan diterapkan sesuai tipe masalah *data mining*.

e. *Evaluation*

Membangun model untuk memastikan model tersebut benar-benar mencapai tujuan bisnis. Pada penelitian ini menggunakan *Root Mean Squared Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE) dalam pengujian tingkat kesalahan.

f. *Deployment*

Penciptaan model umumnya bukan akhir dari proyek, pengetahuan diatur dan disajikan sedemikian rupa sehingga pelanggan dapat menggunakannya. Ini sering melibatkan penerapan model "langsung" dalam proses pengambilan keputusan organisasi, misalnya dalam personalisasi halaman Web secara real-time atau penilaian berulang dari database pemasaran.

3. Hasil dan Pembahasan

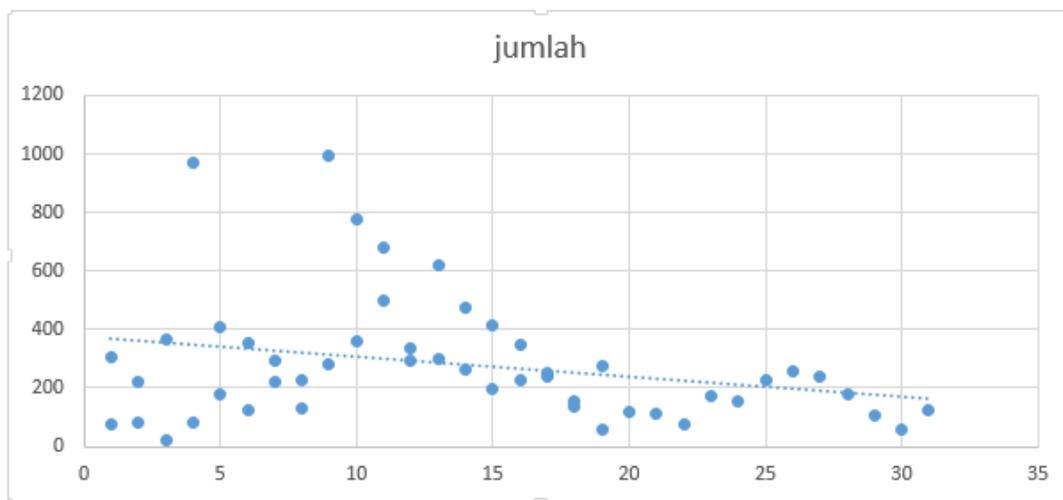
Menjelaskan langkah-langkah yang sesuai metode penelitian. Berikut hasil pembahasan dari penelitian selengkapnya.

a. *Business Understanding*

Dari data yang digunakan, jumlah barang yang terjual setiap harinya tidak selalu stabil. Ada beberapa hari dimana barang sangat terjual banyak di hari tersebut, dan setelah itu penjualan barang menurun drastis untuk beberapa hari. Gambar 2 grafik penjualan barang selama bulan Maret dan April tahun 2022.

b. *Data Understanding*

Data yang telah diperoleh memiliki 50 *file* data dengan 15 atribut di masing-masing file tersebut. Proses selanjutnya adalah melihat dan menentukan apakah dari semua atribut tersebut digunakan atau tidak, karena dari semua atribut tersebut hanya ada beberapa yang akan digunakan nantinya dalam proses *data mining*.



Sumber: Hasil Penelitian (2022)

Gambar 2. Grafik Penjualan

c. *Data Preparation*

Pada tahap *preprocessing data* dilakukan agar nantinya data dapat diolah, proses yang dilakukan adalah penggabungan file data yang telah diperoleh dan membuang beberapa atribut yang nantinya tidak akan digunakan selama proses *data mining*.

No.	bulan	tanggal	jumlah	hpp	pendapatan
1	3	1	308	4017871	4936500
2	3	2	222	2154996	2681950
3	3	3	366	3395289	4142400
4	3	4	81	737581	934200
5	3	5	177	1584034	2000400
6	3	6	123	1185876	1470100
7	3	7	219	1276385	1578350
8	3	8	128	883959	1140600
...
45	4	14	477	4597592	5977150
46	4	15	415	3860535	5058950
47	4	16	348	3042736	3912400
48	4	17	240	1885154	2464000
49	4	18	137	936190	1161000
50	4	19	58	546163	699800

Sumber: Hasil Penelitian (2022)

Gambar 3. Data Penjualan Koperasi Karyawan Usaha Bersama

d. *Modelling*

Selanjutnya adalah menentukan variabel independen dan dependen, serta melakukan perhitungan model regresi. Peneliti menentukan variabel dependen adalah jumlah, dan independen adalah tanggal.

No.	bulan	tanggal (x)	jumlah (y)	XY	X ²
1	3	1	308	308	1
2	3	2	222	444	4
3	3	3	366	1098	9
4	3	4	81	324	16
5	3	5	177	885	25
6	3	6	123	738	36
7	3	7	219	1533	49
8	3	8	128	1024	64
...
45	4	14	477	6678	196
46	4	15	415	6225	225
47	4	16	348	5568	256
48	4	17	240	4080	289
49	4	18	137	2466	324
50	4	19	58	1102	361
		$\Sigma X = 686$	$\Sigma X = 14082$	$\Sigma XY = 169424$	$\Sigma X^2 = 12886$

Sumber: Hasil Penelitian (2022)

Gambar 4. Hasil Data

Kemudian data Gambar 4 diubah ekstensinya menjadi csv agar bisa diolah menggunakan bahasa pemrograman python. Kami melakukan modelling dengan membuat rasio data untuk train dan test sebesar 9:1. Selanjutnya adalah mencari nilai intercept atau konstanta (a) dan slope atau koefisien.

1) Nilai intercept

```
In [605]: lr.intercept_
Out[605]: 376.70535041446874
```

Sumber: Hasil Penelitian (2022)

Gambar 5. Hasil Intercept

Nilai intercept yang didapatkan adalah sebesar 376,7053.

2) Nilai Slope

```
In [604]: lr.coef_
Out[604]: array([-6.78770589])
```

Sumber: Hasil Penelitian (2022)

Gambar 6. Hasil Coefisien

Nilai slope yang didapat adalah sebesar -6,7877.

Berdasarkan hasil perhitungan data, maka didapatkan hasil persamaan regresi linear adalah $Y = -6,7877X + 376,7053$.

3) Prediksi

Setelah mendapatkan hasil persamaan regresi, hal yang selanjutnya dilakukan adalah menghitung hasil prediksi jumlah barang yang akan terjual di masa yang akan datang. Disini peneliti menggunakan data *test* yang diberikan dari hasil perhitungan menggunakan bahasa pemrograman *python*, pada Gambar 7.

```
In [12]: X_test
```

```
Out[12]:
```

tanggal	
2	3
30	31
0	1
15	16
42	12

Sumber: Hasil Penelitian (2022)

Gambar 7. Data Test

$$Y = 375,556 + (-6,845)3$$

$$= 355.021$$

$$Y = 375,556 + (-6,845)31$$

$$= 163.361$$

$$Y = 375,556 + (-6,845)1$$

$$= 368.711$$

$$Y = 375,556 + (-6,845)16$$

$$= 266.036$$

$$Y = 375,556 + (-6,845)12$$

$$= 293.416$$

Dari hasil perhitungan didapatkan prediksi penjualan barang dengan menggunakan regresi linear untuk tanggal 3, 31, 1, 16, dan 12 adalah masing-masing 355.021, 163.361, 368.711, 266.036, 293.416.

e. Evaluation

Selanjutnya adalah melakukan pengujian terhadap model regresi dan melihat seberapa besar kesalahan yang terjadi antara data aktual dan hasil prediksi. Pengujian ini menggunakan RMSE dan MAPE:

1) RMSE

Berikut adalah hasil perhitungan menggunakan RMSE pada rumus 4.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (A_t - F_t)^2}{n}}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^5 (366-355.021)^2 + (125-163.361)^2 + (308-368.711)^2 + (224-266.036)^2 + (336-293.416)^2}{5}}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^5 120.538 + 1471.566 + 3685.826 + 1767.025 + 1813.397}{5}}$$

$$RMSE = 42.091$$

Hasil RMSE yang didapat dari hasil prediksi adalah sebesar 42.091.

2) MAPE

Perhitungan besar kesalahan menggunakan MAPE pada rumus 5.

$$MAPE = \frac{\sum_{t=1}^5 \left| \left(\frac{A_1 - F_1}{A_1} \right) + \left(\frac{A_2 - F_2}{A_2} \right) + \left(\frac{A_3 - F_3}{A_3} \right) + \left(\frac{A_4 - F_4}{A_4} \right) + \left(\frac{A_5 - F_5}{A_5} \right) \right| * 100}{5}$$

$$MAPE = \frac{\sum_{t=1}^5 \left| \left(\frac{366-355.021}{366} \right) + \left(\frac{125-163.361}{125} \right) + \left(\frac{308-368.711}{308} \right) + \left(\frac{224-266.036}{224} \right) + \left(\frac{336-293.416}{336} \right) \right| * 100}{5}$$

$$MAPE = \frac{|0.029 + (-0.306) + (-0.197) + (-0.187) + 0.126| * 100}{5}$$

$$MAPE = 10.7$$

Hasil perhitungan yang didapat menggunakan MAPE adalah sebesar 10.7%.

4. Kesimpulan

Pada penelitian ini melakukan prediksi terhadap jumlah barang yang terjual di Koperasi Karyawan Usaha Bersama. Sebagai kesimpulan dari penelitian tentang hasil penjualan barang ini menggunakan metode regresi linear, hasil prediksi mendapatkan *error* sebesar 42,091 pada RMSE dan 10,7% pada MAPE sehingga pengujian MAPE lebih baik jika dibandingkan dengan RMSE

Daftar Pustaka

- Ayuni, G. N., & Fitriana, D. (2019). Penerapan Metode Regresi Linear untuk Prediksi Penjualan Properti pada PT XYZ. *Jurnal Telematika*, 14(2), 79–86.
- Gustriansyah, R. (2017). Analisis Metode Single Exponential Smoothing Dengan Brown Exponential Smoothing Pada Studi Kasus Memprediksi Kuantitas Penjualan Produk Farmasidi Apotek. *Seminar Nasional Teknologi Informasi Dan Multimedia 2017*, 3, 5–12.
- Hakimah, M., Muhima, R. R., & Yustina, A. (2015). Persediaan Barang dengan Metode Trend Projection. *SimanteC*, 5(1), 37–48.
- Indarwati, T., Irawati, T., & Rimawati, E. (2019). Penggunaan Metode Linear Regression Untuk Prediksi Penjualan Smartphone. *Jurnal Teknologi Informasi Dan Komunikasi (TIKOMSiN)*, 6(2). <https://doi.org/10.30646/tikomsin.v6i2.369>
- Katumba, P., & Djoh, R. K. (2017). *Prediksi Tingkat Produksi Kopi Menggunakan Regresi Linear*.
- Larose, D. T. (2005). An Introduction to *Data mining*. In *Structure and Bonding* (Vol. 134). https://doi.org/10.1007/430_2009_1

- Nafi'iyah, N., & Rakhmawati, E. (2021). *Analisis Regresi Linear dan Moving Average Dalam Memprediksi Data Penjualan Supermarket* (Vol. 12, Issue 1).
- Ong, A. S. H. (2009). Penerapan Metode Regresi Linear Sederhana untuk Prediksi Harga Beras di Kota Padang. *Journal of Science and Technology in the Tropics*, 5(1), 3.
- Pakaja, F., & Naba, A. (2018). Peramalan Penjualan Mobil Menggunakan Jaringan Syaraf Tiruan dan Certainty Factor. *Neural Networks*, 6(1), 23–28.
- Peraturan Pemerintah Republik Indonesia. (2012). *Undang-Undang Nomor 17 Tahun 2012 Tentang Perkoperasian*.
- Prasetyowati, E. (2018). Aplikasi Penentuan Harga Pokok Produksi Batik Madura Dengan Metode Activity Based Costing dan Analisis Regresi Linier. *JUTI: Jurnal Ilmiah Teknologi Informasi*, 16(1), 48. <https://doi.org/10.12962/j24068535.v16i1.a690>
- Syafuruddin, M., Hakim, L., & Despa, D. (2014). Metode Regresi Linier untuk Prediksi Kebutuhan Energi Listrik Jangka Panjang (Studi Kasus Provinsi Lampung). *Jurnal Informatika Dan Teknik Elektro Terapan*, 2(2). <https://doi.org/10.23960/jitet.v2i2.237>