JURNAL INFOTEL
Informatics - Telecommunication - Electronics
Website: http://ejournal.st3telkom.ac.id/index.php/infotel
ISSN: 2085-3688; e-ISSN: 2460-0997

Implementation of multiple linear regression to estimate profit on sales of screen printing equipment

Khairul[1], Asyhari Hadi Nasyuha[2,*], Ali Ikhwan[3], Moustafa H. Aly[4], Ahyanuardi[5]
[1]Faculty of Science and Technology, Universitas Pembangunan Panca Budi
[2]Faculty of Technology Information, Universitas Teknologi Digital Indonesia
[3]Faculty Sains Technology, Universitas Islam Negeri Sumatera Utara
[4]Arab Academy for Science, Technology and Maritime Transport
[5]Faculty of Engineering, Universitas Negeri Padang
[1]Jl. Jendral Gatot Subroto, KM. 4,5, Sei Sikambing, Medan, Indonesia
[2]Jl. Raya Janti Karang Jambe, No. 143, Yogyakarta 55198, Indonesia
[3]Jl. Willem Iskandar, Pasar V, Medan Estate, Deli Serdang, Sumatera Utara 20371, Indonesia
[4]32J9+M5V, Al Giza Desert, Giza Governorate 3650111, Egypt
[5]Jl. Prof. Dr. Hamka, Air Tawar Padang, Sumatera Barat, Indonesia
[*]Corresponding email: asyahrihadi@gmail.com

Received 24 March 2023, Revised 14 April 2023, Accepted 6 June 2023

Abstract — Traditional marketing strategies are no longer practical to implement because the process requires more costs and longer time to disseminate information. Marketing strategy is one of the key aspects of business success, be it for large or small companies. An effective marketing strategy must be able to attract consumer attention, build brand awareness, increase sales, and strengthen the company's market position. In developing a marketing strategy, it is necessary to carry out market and competitor analysis, determine market targets, determine product or service positioning, choose the right distribution channel, determine the appropriate price, and develop an effective promotional program. In the digital era, marketing strategies must also include the use of computer technology. Data mining is a science that discusses knowledge from previous data to estimate the amount of production in the future. Data mining is a term used to find hidden knowledge in databases. "Data mining is a semi-automated process using statistics, mathematics, artificial intelligence and machine learning techniques to extract and identify valuable and useful information in large databases. For this reason, it is necessary to solve this problem by using one of the five methods in the field of Data Mining, namely the multiple linear regression method, where this method will analyze the influential variables and can make estimates. Multiple linear regression is a method that can be used to analyze data and obtain meaningful conclusions about a relationship between one variable and another. This relationship is generally expressed by a mathematical equation which states the relationship between the independent and dependent variables in the form of a simple equation.

Keywords – data mining, multiple linear regression, sales profit

I. Introduction

Screen printing is printing ink on the top of the t-shirt media with tools such as screens or screen printing film [1]. At this time, the telang t-shirt screen printing service has developed into a particular type of business that is quite promising because there are more and more orders for t-shirt screen printing, ranging from just merchandise, the interests of an event, for community uniforms, to campaign purposes. The screen-printing business is increasingly growing and expanding day by day. This business is increasingly favored by the community, especially among young people. Various types of screen-printing businesses, screen printing, is a printing process technique that uses a cloth (screen) with a certain level of pores and is usually made of Nylon or Silk. After that, this screen is given an image according to the designer's wishes, previously designed on HVS or tracing paper. This fabric is firmly attached to a rectangular piece of wood to produce a flat and neat screen and printout. After

being given photo resistance and irradiating, it must be slowly and carefully watered so that the pattern is visible, and then parts that the ink can pass through according to the picture made earlier. The process is by pouring ink on the screen's surface, and after, it is smeared using a board or rack made of rubber. One screen can only be used to make one color, while making color variants in one design requires a unique tool for precision.

Traditional marketing strategies are no longer practical to implement because the process requires more costs and time to disseminate much longer information [2]. In this modern era, many advanced technologies can shift traditional marketing into modern marketing, namely by turning into a marketing model by utilizing digital media. Digital media such as blogs, Instagram, Facebook, and so on have proliferated. Entrepreneurs can upload photos of screen-printing designs, price lists, and produced works that attract customers' interest.

II. Research Method

This section discusses data mining and multiple linear regression.

A. Data Mining

The terms data mining and knowledge discovery in databases are often used interchangeably to describe the process of extracting information still hidden in a large database [3]–[5]. One of the stages in the whole KDD process is Data Mining [6], [7]. In the data mining process, which is usually called knowledge discovery in databases. Knowledge discovery in databases is the application of scientific methods to a data mining [8], [9].

Data mining is a step to finding interesting patterns and knowledge from large amounts of data [7], [10]. Data mining is a search and analysis of large amounts of data and aims to find the meaning of patterns and rules. Data mining is a step of extracting or extracting data that was unknown before but can be understood and valuable from large databases [8], [11]. It is used to make essential business decisions [12], [13]. From the three responses from the experts above, an outline can be drawn; data mining is a step to find and analyze a data collection (database) in order to find an interesting pattern in order to extract information and knowledge accurately and potentially [11], [14], [15].

B. Multiple Linear Regression

Multiple linear regression method is a technique in analyzing data by trying and looking for a relationship between two or more variables [16], [17], especially between variables that contain cause and effect [18]–[20]. Multiple Linear Regression Analysis is often used as a preference analysis to determine the effect of one variable or variable with other variables [21]–[23].The general formula used in regression analysis is written in (1) [24]–[26].

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

where Y is a dependent variable, X is a independent variable, a is a constant variable, and b is a linear regression disrection coefficient.

To forecast the influence of a predictor variable on a criterion variable or to demonstrate the presence or absence of a functional connection between an independent variable (X) and a dependent variable.

The multiple linear regression model is a model that describes and explains the relationship of one dependent variable to two or more predictor variables [23], [26], [27]. Multiple linear regression is an extension of simple linear regression. The expansion can be seen from the number of independent variables in the regression model [18], [28]–[30]. The general form of multiple linear regression can be stated statistically as stated in (2) [31].

$$Y = a + b1X1 + b2X2 + bnXn \tag{2}$$

where Y is bound variable, X is independent variable, a, $b1$, $b2$, bn are regression parameters/regression coefficient.

III. Result

The research methodology is a process or scientific method. In the first stage, an analysis of the problem of estimating profits from the sale of screen printing equipment is carried out. The next stage is to collect data related to sales and the equipment sold. Then apply calculations with multiple linear regression analysis to obtain the estimated profit obtained. The following are the variables used as benchmarks in estimating sales results. It can be seen in Table 1 related to the variable data used.

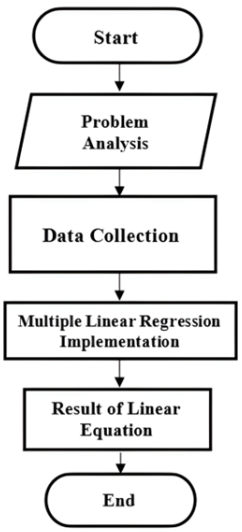

Fig. 1. Flowchart research process.

Table 1. Name of Variables Data

No.	Variable	Name of Variable	Description
1.	$X1$	Cost of goods	The initial price of the purchase of goods
2.	$X2$	Selling price	The retail price of each item
3.	$X3$	Total shopping	The amount of spending when making a purchase
4.	Y	Total sales	Revenue earned from sales

Table 2. Equipment Sales Revenue Data

Period	Cost of Goods Sold ($X1$)	Selling Price ($X2$)	Total Shopping ($X3$)	Total Sales (Y)
Jan 2020 – Feb 2020	Rp. 839,000	Rp. 1,115,000	Rp. 158,775,000	Rp. 216,400,000
Feb 2020 – Mar 2020	Rp. 719,000	Rp. 980,000	Rp. 185,075,000	Rp. 252,000,000
Mar 2020 – Apr 2020	Rp. 859,000	Rp. 1,138,000	Rp. 173,575,000	Rp. 235,940,000
Apr 2020 – Mei 2020	Rp. 859,000	Rp. 1,138,000	Rp. 201,750,000	Rp. 276,340,000
May 2020 – Jun 2020	Rp. 749,000	Rp. 1,035,000	Rp. 114,915,000	Rp. 156,450,000
Jun 2020 – Jul 2020	Rp. 989,000	Rp. 129,000	Rp. 122,320,000	Rp. 167,150,000
Jul 2020 – Aug 2020	Rp. 884,000	Rp. 117,300	Rp. 130,200,000	Rp. 179,540,000
Aug 2020 – Sept 2020	Rp. 739,000	Rp. 1,015,000	Rp. 188,025,000	Rp. 256,600,000
Sep 2020 – Oct 2020	Rp. 869,000	Rp. 1,165,000	Rp. 190,475,000	Rp. 260,350,000
Oct 2020 – Nov 2020	Rp. 869,000	Rp. 1,110,000	Rp. 225,150,000	Rp. 259,100,000
Nov 2020 – Des 2020	Rp. 914,000	Rp. 1,188,000	Rp. 219,700,000	Rp. 291,240,000
Des 2020 – Jan 2021	Rp. 1,064,000	Rp. 1,423,000	Rp. 187,650,000	Rp. 261,790,000

The steps in this research are analyzing the problem of estimating profits from printing equipment sales, then collecting data from sales, inventory to sales profits each month, after that calculations will be carried out with multiple linear regression implementation, and after that get the results of this research.

This data collection was carried out to obtain data related to direct review of the sales results of the screen-printing equipment each month. The following is sales revenue data for equipment which can be seen in Table 2.

It takes an algorithm that will be used as a step in an analysis in processing and solving a problem. Fig. 2 is a flowchart or flow of problem solving using the Multiple Linear Regression method in this study.

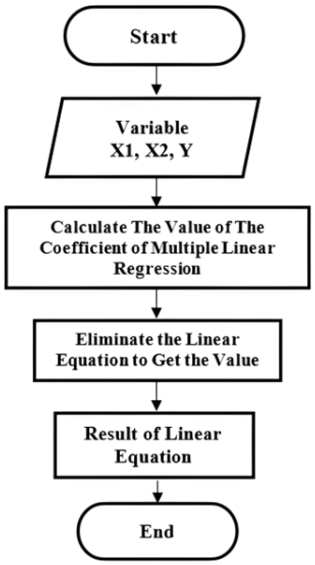

Fig. 2. Multiple linear regression flowchart.

In Fig. 2, the steps in calculating multiple linear regression, first, determine the variables $X1$, $X2$, Y. After that, calculate the value of the multiple linear regression coefficients of the variables that have been determined, then eliminate the linear equation to get the final value. Furthermore, conclude the final result of the calculation of linear equations.

A. Calculate the Value of the Coeffecient of Multipple Linear Regression

The regression coefficient is vital in analyzing regression. The benefit of the regression coefficient is to form a model of the regression equation on a problem under study. Calculating the regression coefficient can be done by calculating the mathematical equation from the equation formula pattern that has been determined in the provisions of multiple linear regression. As can be seen in Table 3 below about the regression coefficient.

From the Table 3, it is known the value:

$\sum X1$ = 103.53

$\sum X2$ = 137.7

$\sum X3$ = 2,097.61

$\sum Y$ = 2,812.9

$\sum X12$ = 904.116

$\sum X22$ = 1,596.079

$\sum X32$ = 381,220.540

$\sum X1X2$ = 1,201.011

$\sum X1X3$ = 18,114.219

$\sum X2X3$ = 24,056.767

$\sum X1Y$ = 24,297.6925

$\sum X2Y$ = 32,292.4835

$\sum X3Y$ = 508,787.106

Table 3. Regression Coefficient

No.	$X1$	$X2$	$X3$	Y	$X1*Y$	$X2*Y$	$X3*Y$
1.	8.39	11.15	158.775	216.4	1,815.596	2,412.86	34,358.91
2.	7.19	9.8	185.075	252	1,811.88	2,469.6	46,638.9
3.	8.59	11.38	173.575	235.94	2,026.7246	2,684.9972	40,953.2855
4.	8.59	11.38	201.75	276.34	2,373.7606	3,144.7492	55,751.595
5.	7.49	10.35	114.915	156.45	1,171.8105	1,619.2575	17,978.45175
6.	9.89	12.9	122.32	167.15	1,653.1135	2,156.235	20,445.788
7.	8.84	11.73	130.2	179.54	1,587.1336	2,106.0042	23,376.108
8.	7.39	10.15	188.025	256.6	1,896.274	2,604.49	48,247.215
9.	8.69	11.65	190.475	260.35	2,262.4415	3,033.0775	49,590.16625
10.	8.69	11.1	225.15	259.1	2,251.579	2,876.01	58,336.365
11.	9.14	11.88	219.7	291.24	2,661.9336	3,459.9312	63,985.428
12.	10.64	14.23	187.65	261.79	2,785.4456	3,725.2717	49,124.8935
Total	103.53	137.7	2097.61	2,812.9	24,297.6925	32,292.4835	508,787.106

From the calculation of the coefficient of multiple linear regression, then making a linear equation.

1) First Process

$$\sum Y = na + b1 \sum X1 \tag{3}$$

Eq. (3) and the regression coefficient, the result is $2,812.9 = 12a + 103.53b1 + 137.7b2 + 2,097.61b3$.

2) Second Process

$$\begin{aligned}\sum X1Y = \\ a\sum X1 + b1\sum X12+ \\ b2\sum X1X2 + b3\sum X1X3\end{aligned} \tag{4}$$

Eq. (4) and the regression coefficient, the result is $24,297.6925 = 103.53a + 904.116b1 + 1,201.011b2 + 18,114.219b3$.

3) Third Process

$$\begin{aligned}\sum X2Y = \\ a\sum X2 + b1\sum X1X2+ \\ b2\sum X22 + b3\sum X2X3\end{aligned} \tag{5}$$

Eq. (5) and the regression coefficient, the result is $32,292.4835 = 137.7a + 1,201.011b1 + 1,596.079b2 + 24,056.767b3$.

4) Fourth Process

$$\begin{aligned}\sum X3Y = \\ a\sum X3 + b1\sum X1X3+ \\ b2\sum X2X3 + b3\sum X32\end{aligned} \tag{6}$$

Eq. (6) and the regression coefficient, the results is $508,787.106 = 2,097.61a + 18,114.219b1 + 24,056.767b2 + 381,220.540b3$.

5) Fifth Process

Performing the elimination process from (3):

$$\begin{aligned}2,812.9 =& 12a + 103.5b1 + 137.7b2+ \\ & 2,097.61b3 \cdots (3) \times 103,53 \\ 291,219.537 =& 1,242.36a + 10,718.4609b1+ \\ & 14,256.081b2 + 217,165.5633b3\end{aligned}$$

Also performing the elimination from (4) as follows:

$$\begin{aligned}24,297.6925 =& 103.53a + 904.116b1 + 1,201.011b2+ \\ & 18,114.219b3 \cdots (4) \times 12 \\ 291,572,31 =& 1,242.36a + 10,849.392b1+ \\ & 14,412.132b2 + 217,370.628b3\end{aligned}$$

Therefore, $-352.773 = -130.9311b1 + (-156.051b2) + (-205.0647b3)$.

6) Sixth Process

Performing the elimination process from (3):

$$\begin{aligned}2,812.9 =& 12a + 103.53b1 + 137.7b2+ \\ & 2,097.61b3 \cdots (3) \times 137.7 \\ 387,336.33 =& 1,652.4a + 14,256.081b1+ \\ & 18,961.29b2 + 288,840.897b3\end{aligned}$$

Also performing the elimination from (5) as follows:

$$\begin{aligned}32,292.4835 =& 137.7a + 1,201.011b1 + 1,596.079b2 \\ & + 24,056.767b3 \cdots (5) \times 12 \\ 387,509.802 =& 1,652.4a + 14,412.132b1+ \\ & 19,152.948b2 + 288,681.204b3\end{aligned}$$

Therfore, $173.472 = -156.051b1 + (-191.658b2) + 159,693b3$.

7) Seventh Process

Performing the elimination process from (3):

$$\begin{aligned}2,812.9 =& 12a + 103.53b1 + 137.7b2+ \\ & 2,097.61b3 \cdots (3) \times 2,097.61 \\ 5,900,367.169 =& 25,171.32a + 217,165.5633b1+ \\ & 288,840.897b2 + 4,399,967.712b3\end{aligned}$$

Also performing the elimination from (4) as follows:

$$\begin{aligned}508,787.106 =& 2,097.61a + 18,114.219b1+ \\ & 24,056.767b2 + 381,220.540b3 \\ 6,105,445.272 =& 25,171.32a + 2,173,370.628b1+ \\ & 288,681.204b2 + 4,574,646.48b3\end{aligned}$$

Therefore, $-205,078.103 = -205.0647b1 + 159.693b2 + (-174,678.7679b3)$.

8) Eighth Process

Performing the elimination process from *Fifth Process*:

$$\begin{aligned} -352.773 = &-130.9311b1 + -156.051b2 + \\ &-205.0647b3 \cdots \times (-156.051) \\ 55,050.735 = &20,431.9291b1 + 24,351.9146b2 + \\ &32,000.551b3 \end{aligned}$$

Also performing the elimination from *Sixth Process* as follows:

$$\begin{aligned} -173.472 = &-156.051b1 + (-191.658b2) + \\ &159.693b \cdots \times (-130.9311) \\ 22,712 = &20,431.9291b1 + 25,093.99276b2 + \\ &(-20,908.78b3) \end{aligned}$$

Therefore, $-26,778,760.12 = 52,909.33165b2 + (-22,828,831.7b3)$.

9) Ninth Process

Performing the elimination process from *Fifth Process* as follows:

$$\begin{aligned} -352.773 = &-130.9311b1 + (-156.051b2)(10) + \\ &(-205.0647b3) \times (-1,956,205.06) \\ 72,341.49448 = &26,849.34674b1 + 32,000.5515b2 + \\ &42,051.53119b3 \end{aligned}$$

Also performing the elimination from *Seventh Process* as follows:

$$\begin{aligned} -205,078.103 = &-205.0647b1 + 159.693b2 + \\ &(-174,678.7679b3) \cdots \times (-130.93) \\ 26,851,101.61 = &26,849.34674b1 + (-20,908.78015) \\ &+ 22,870,883.23 \end{aligned}$$

Therefore, $-26,778,760.12 = 52,909,33165b2 + (-22,828,831.7b3)$.

10) Tenth Process

Next, perform the elimination process between equation (8) and equation (9) as follows:

$$\begin{aligned} 32,337.85569 = &-742.07816b2 + 52,909.331b3 \\ &\cdots \times 52,909.33165 \\ 1,710,974,332 = &-39,262,859.63b2 + \\ &2,799,397,376b3 \end{aligned}$$

$$\begin{aligned} -26,778,760.12 = &-52,909.331b2 + \\ &(-22,828,831.7b3) \\ &\cdots \times (-742.0781628) \\ 19,871,933,112 = &-39,262,859.63b2 + \\ &(-16,940,777,487b3) \end{aligned}$$

Therefore, $-18,160,958,780 = -14,141,380,111b3$.

$$\begin{aligned} b3 &= \frac{-18,160,958,780}{-14,141,380,111} \\ &= 1.284242318 \end{aligned}$$

By the result of $b3$, therefore $b2$ can be calculated as folows:

$$\begin{aligned} 32,337.85569 = &-741.078162 8b2 + 52,909.33165 \\ &(1.284242318) \\ 32,337.85569 = &-741.078162 8b2 + 67,948.40272 \\ -35,610.54703 = &-741.078162 8b2 \\ 32,337.85569 = &-741.078162 8b2 \end{aligned}$$

$$\begin{aligned} b2 &= \frac{32,337.85569}{-741.078162 8b2} \\ &= -43.63622802 \end{aligned}$$

Case testing based on the equation:

It is assumed that the cost of goods is Rp. 839,000, Selling Price is Rp. 1,115,000, and Total Expenditure is Rp. 158,775,000. By using data normalization, Cost of goods = Rp. 839,000 = 8.39; Selling Price = Rp. 1,115,000 = 11.15; and Total Expenditure = Rp.158,775,000 = 158,775.

By using (2), therefore:

$$\begin{aligned} Y = &56.05531667 – (-52.69105853)X1 + \\ &(-43.63622802)X2 + 1.284242318X3 \\ = &56.05531667 – (-52.69105853)(8.39) + \\ &(-43.63622802)(11.15) + 1.284242318(158.775) \\ = &56.05531667 + 442.0779811 - 48.65439424 + \\ &203.905574 \\ = &Rp.215.494292 \times 1,000,000 \\ = &Rp.215,494,292 \end{aligned}$$

Based on the above calculations, it can be obtained that the value of the sales obtained by the Cahaya Sablon Shop is Rp. 215,494,292. Analysis of the equations resulting from calculations 1 to 10 are the results of calculations that are interconnected in each result, which is helpful for seeing between the independent variable and the dependent variable and being able to predict the value of the dependent variable when the independent variables have a known final result value.

IV. Discussion

Based on this research shows success in analyzing the sale of screen printing equipment at the Cahaya Sablon Shop using Multiple Linear Regression based on the data obtained. At the same time, applying the Multiple Linear Regression method can estimate sales results at the Cahaya Sablon Store as a desktop application. The next stage is to design and build a system to help the Cahaya Sablon Store estimate sales

results using a data mining technique that adopts the Multiple Linear Regression method. Multiple Linear Regression Is a method used to analyze data and obtain meaningful conclusions about a variable's relationship to other variables. There are still deficiencies that can be developed in this study, namely by adding another testing technique that is contained in data mining.

V. Conclusion

Based on the analysis of the problems that occur in the research on estimating the sale of screen printing tools using the multiple linear regression method as a type of statistical analysis, to predict the value of the dependent variable when the independent variables have a known value, the following conclusions can be drawn: Cost of Goods, Selling Price, Total Expenses, and Total Sales are used as variables that affect Total Sales, which will then be calculated for Estimation. Based on the calculation, the selling value of the Cahaya Sablon store is Rp. 215,494,292.

Acknowledgment

We thank those who have assisted in this research process.

References

[1] L. Jahangir, "Different methods of t shirt printing: A comparative analysis," *Canadian Journal of Researcher's Society*, vol. 8, no. 1, 2018.

[2] E. Pantano, C.-V. Priporas, and G. Migliano, "Reshaping traditional marketing mix to include social media participation," *Eur. Bus. Rev.*, vol. 31, no. 2, pp. 162–178, 2019, doi: 10.1108/EBR-08-2017-0152.

[3] R. Patil and S. Tamane, "A comparative analysis on the evaluation of classification algorithms in the prediction of diabetes," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, pp. 3966–3975, 2018, doi: 10.11591/ijece.v8i5.pp3966-3975.

[4] W. Wen, Z. Yuan, and Y. Yuan, "Improving RETECS method using FP-Growth in continuous integration," in *2018 5th IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS)*, 23-25 November 2018, Nanjing, China. doi: 10.1109/CCIS.2018.8691385.

[5] Y. Miao, J. Lin, and N. Xu, "An improved parallel FP-growth algorithm based on Spark and its application," in *Chinese Control Conference (CCC)*, 27-30 July 2019, Guangzhou, China. doi: 10.23919/ChiCC.2019.8866373.

[6] A. J. P. Sibarani, "Implementasi data mining menggunakan algoritma apriori untuk meningkatkan pola penjualan obat," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 2, pp. 262–276, 2020, doi: 10.35957/jatisi.v7i2.195.

[7] S. Bachhety, R. Singhal, and R. Jain, "Intelligent data analysis with data mining: Theory and applications," in *Intelligent Data Analysis: From Data Gathering to Data Comprehension*, 2020.

[8] M. Z. Hossain, M. N. Akhtar, R. B. Ahmad, and M. Rahman, "A dynamic K-means clustering for data mining," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 13, no. 2, pp. 521–526, 2019, doi: 10.11591/ijeecs.v13.i2.pp521-526.

[9] A. H. Nasyuha, Z. Zulham, and I. Rusydi, "Implementation of K-means algorithm in data analysis," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 20, no. 2, pp. 307-313, 2022, doi: 10.12928/telkomnika.v20i2.21986.

[10] A. Penira and A. Ikhwan, "Data mining memprediksi kecenderungan calon nasabah dalam memilih produk asuransi PT. AJS Bumiputera Medan," *Jurnal Ilmu-ilmu Informatika dan Manajemen STMIK*, vol. 16, no. 1, pp. 61–70, 2022.

[11] X. Ma, Y. Li, R. Liu, Y. Zhang, L. Ma, and Z. Gao, "Frequent itemsets mining of SCADA data based on FP-growth algorithm," in *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*, 30 October 2020 - 01 November 2020, Wuhan, China. doi: 10.1109/EI250167.2020.9346885.

[12] Y. T. S. Carolino and S. R. Nunes, "The maturity and efficiency of it governance processes based on cobit 5: A case of a health sector organization in Portugal," in *Iberian Conference on Information Systems and Technologies, CISTI*, 19-22 June 2019, Coimbra, Portugal. doi: 10.23919/CISTI.2019.8760913.

[13] H. He and G. X. Li, "Strategies for building a complete and efficient online shopping supply chain system," in *E3S Web of Conferences*, 2021, vol. 253. doi: 10.1051/e3sconf/202125303030.

[14] Y. Yang and T. Chen, "Analysis and visualization implementation of medical big data resource sharing mechanism based on deep learning," *IEEE Access*, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2949879.

[15] T. R. S. Mary and S. Sebastian, "Predicting heart ailment in patients with varying number of features using data mining techniques," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 4, pp. 2675-2681, 2019, doi: 10.11591/ijece.v9i4.pp2675-2681.

[16] J. Qidi, "Research on influencing factors of retail sales in e-commerce market," in *2021 2nd International Conference on E-Commerce and Internet Technology (ECIT)*, 05-07 March 2021, Hangzhou, China. doi: 10.1109/ECIT52743.2021.00011.

[17] S. I. Bangdiwala, "Regression: multiple linear," *International Journal of Injury Control and Safety Promotion*, vol. 25, no. 2. pp. 232–236, 2018. doi: 10.1080/17457300.2018.1452336.

[18] V. C. Georgopoulos, A. T. Rajappa, K. Soriano, C. Zeimer, and G. Malandraki, "Feasibility of a neural network in predicting aspiration severity post supratentorial stroke," *Dysphagia*, vol. 34, no. 6, 2019.

[19] F. Khademi, M. Akbari, S. M. Jamal, and M. Nikoo, "Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete," *Front. Struct. Civ. Eng.*, vol. 11, no. 1, pp. 90–99, 2017, doi: 10.1007/s11709-016-0363-9.

[20] X. Huang, H. Wang, W. Luo, S. Xue, F. Hayat, and Z. Gao, "Prediction of loquat soluble solids and titratable acid content using fruit mineral elements by artificial neural network and multiple linear regression," *Scientia Horticulturae*, vol. 278, 2021,.doi: 10.1016/j.scienta.2020.109873.

[21] L. Yang, "Research on quantitative evaluation method of teachers based on multiple linear regression," in *2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, 16-17 January 2021. Beihai, China. doi: 10.1109/ICMTMA52658.2021.00196.

[22] Y. Yuan, Y. Li, and H. Bao, "An empirical study on the kane-change duration of naturalistic driving based on multiple linear regression model," 2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS), 20-22 March 2020, Dalian, China. doi: 10.1109/ICAIIS49377.2020.9194893.

[23] P. Huang, P. Gu, X. Cao, K. Jiang, B. Duan, and C. Zhang, "Battery state of health estimation based on multiple linear regression by particle swarm optimization," in *2021 40th Chinese Control Conference (CCC)*, 26-28 July 2021, Shanghai, China. doi: 10.23919/CCC52363.2021.9549690.

[24] O. Chantarakasemchit, S. Nuchitprasitchai, and Y. Nilsiam, "Forex rates prediction on EUR/USD with simple moving average technique and financial factors," in *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 24-27 June 2020, Phuket, Thailand. doi: 10.1109/ECTI-CON49241.2020.9157907.

[25] S. Wong, "Stock price prediction model based on the short-term trending of KNN method," in *2020 7th International Conference on Information Science and Control Engineering (ICISCE)*, 18-20 December 2020, Changsha, China. doi: 10.1109/ICISCE50968.2020.00273.

[26] D. Alita, A. D. Putra, and D. Darwis, "Analysis of classic assumption test and multiple linear regression coefficient test for employee structural office recommendation," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 3, pp. 295 - 306, 2021, doi: 10.22146/ijccs.65586.

[27] A. K. Kadam, V. M. Wagh, A. A. Muley, B. N. Umrikar, and R. N. Sankhua, "Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India," *Model. Earth Syst. Environ.*, vol. 5, no. 3, pp. 951–962, Sep. 2019, doi: 10.1007/s40808-019-00581-3.

[28] Z. Zhong, T. R. Carr, X. Wu, and G. Wang, "Application of a convolutional neural network in permeability prediction: A case study in the Jacksonburg-Stringtown oil field, West Virginia, USA," *Geophysics*, vol. 84, no. 6, 2019, doi: 10.1190/geo2018-0588.1.

[29] S. Feng, C. L. Philip Chen, L. Xu, and Z. Liu, "On the accuracy-complexity tradeoff of fuzzy broad learning system," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 10, 2021, doi: 10.1109/TFUZZ.2020.3009757.

[30] S. Dagdanpurev, G. Sun, L. Choimaa, S. Abe, and T. Matsui, "Clinical application of multiple vital signs-based infection screening system in a Mongolian hospital: Optimization of facial temperature measurement by thermography at various ambient temperature conditions using linear regression analysis," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2018, vol. 2018-July. doi: 10.1109/EMBC.2018.8513513.

[31] F. Farjana, S. K. Singha, and A. Al Farabi, "Cuffless blood pressure determination using photoplethysmogram (PPG) signal based on multiple linear regression analysis," in *2021 International Conference on Science & Contemporary Technologies (ICSCT)*, 05-07 August 2021, Dhaka, Bangladesh. doi: 10.1109/ICSCT53883.2021.9642626.