ModellingUnder Uncertainty Business Competitive With Robust Estimation Measurement Model

Marischa Elveny*^a, Rahmad Syah^b, Al-khowarizmi^c

^a*Department of Information Technology, Universitas Sumatera Utara, Medan, Indonesia.

^b Department of Informatics, Universitas Medan Area, Medan, Indonesia

^c Department Information System, Universitas Muhammadiyah Sumatera Utara, Medan Indonesia.

(*Corresponding author's e-mail: marischaelveny@usu.ac.id)

Article History: Received: 10 November 2020; Revised 12 January 2021 Accepted: 27 January 2021; Published online: 5 April 2021

Abstract: Business insight can be characterized as a bunch of procedures and apparatuses for the obtaining and change of crude information into significant and valuable data for business examination purposes. Information requires strategies and programs that competing organizations in which e-metrics cannot ignore it. Electronic-Metrics or commonly known as emetrics are data created based on electronic-based customer behavior (e-customer behavior). Extracting information on a competitive organization involves a similarity that is used to find a connection or linkage between organizational behavior. Various new ways are used to achieve success, one of which is with an electronic-based business, but with such a large number of variations, the uncertainty in business is also increasingly difficult to predict. Especially in predicting what activities will often occur in the next few years. Prediction is a process of systematically estimating something that is most likely to happen in the future based on past and present information. To keep up with the development of the company, it is necessary to optimize the metrics for the business. The purpose of optimization is to find the minimum or maximum value of a problem, whether the value of a company produces the desired results. Optimization can be done with robust nonparametric regression, where robust regression is used to detect outliers and provide results that are resistant to outliers. MM (Method of Moment) estimation is one of the estimates in robust regression that can be used for detected outlier data on the independent and dependent variables. Based on the parameter significance test, it was found that merchants who sell food are more profitable than other merchants. With a difference in iteration value 23.9 and a comparison of iteration value 53.1. Based on the parameter significance test conducted, it was found that merchants who sell food are more.

Keywords: Business Intelligence, Optimization, Prediction, Business Metrics, Robust, Nonparametric Regression, Robust Estimation-MM

1. Introduction

Business Intelligence and Analytics (BIA) is an assortment of techniques, measures, designs, applications and innovations that gather and change crude information into significant and valuable data used to empower more compelling methodologies, strategic experiences, and tasks in dynamic [1]. Business knowledge (BI) can be depicted as a bunch of methods and instruments for the securing and change of crude information into significant and valuable data for business investigation purposes. BI can be used to support a variety of business decisions from operational to strategic. Basic operating decisions include determining the position or price of the product. Strategic business decisions include priorities, goals, and direction at the broadest level [2]. Information is not only an important factor for knowledge acquisition but is a determining factor for the profit of each scenario in competition and success. Information requires strategies and programs that cannot be ignored by competitive organizations where e-metrics play a role [3]. Electronic-Metrics or purported e-measurements are information made dependent on electronic-based client conduct (e-client behavior) [4]. The advanced time has carried exhaustive changes to the network, this can be seen in introducing business plans utilizing e-measurements that give accommodation on how (clients) carry on electronically and give data that can be applied [5]. Emeasurements empower the economy to become quicker, and better, on the grounds that in e-measurements the information acquired must be legitimate for all individuals from the association. Likewise, dispersed emeasurements reflect pieces of the pie and markets that are changing quickly so they should adjust to changes.

Extracting information on a competitive organization involves a similarity that is used to find a connection or linkage between competitive organizational behavior. The comparability of two articles is estimated dependent on the idea of nearness, which is utilized to give a limited worth. There are numerous likeness measures, however including the request for the segments of the article requires an extraordinary methodology [6]. The approach is carried out by analyzing customer profiles based on the level of behavior similarity, Basic measurements of

distances, attributes, times, places, and transactions between social actors are used to trace the origin of the information, or can be used to predict social behaviour [7].

Optimization of prediction is needed as a process of systematically estimating something that is most likely to happen in the future based on past and present information [8] [9]. From this relationship can be made a model that can predict what will happen in the future [10] [11]. In the nonparametric regression model, it is possible that there are outliers (outliers) which cause some regression assumptions not to be fulfilled so that the predictive value is less accurate, therefore a method that can handle outlier data is used, namely the robust regression method [12] [13].

Robust is one of the displaying techniques for handling information enhancement issues that are dubious and just known to have a place with some arrangement of vulnerabilities, within the sight of exceptions [14]. Robust will probably locate an attainable ideal or close to ideal answer for every conceivable acknowledgment of a questionable situation [15]. In accordance with the profile similarity profit optimization model and looking for outliers several references were taken.

According to previous research [16], introducing appropriate metrics to measure process similarity based on behavior profiles. [17] Robust CMARS provides a solution to the problem by optimizing the sensitivity to disturbances in problem parameters. [18] Presents the M estimate, S estimate and MM estimate in robust regression to determine the regression model. M estimation is an extension of the maximum likelihood method and is where the estimated value is not much influenced by small changes in data [19].

Based on the research that has been done before, this research was conducted to can produce an optimization model in predicting profits with e-metrics data that has not been done by previous research, based on customer profiles by approaching robust nonparametric regression. Where robust is used to find outliers that occur in competitive organizations, based on several variables that will be used so that this modeling is expected to be a model in predicting profits, which will later be used as decision making for competitive organizations.

2. Research Purpose

One common wish across all industries is to be able to quickly and easily analyze financial data in a database in order to make smarter decisions about the future. Transactional data is good for tracking what is happening in an organization, but not suitable for finding out why something happens or predicting future performance [18]. In other words, it is time to use tools to extract data from the same system and organize it in a useful way to support the decision-making process.

Information requires strategies, methods and tools that cannot be ignored by competitive organizations through e-metrics data [19]. On the one hand, not all customer profiles are similar to one another, there is the possibility of outlier for every transaction. On the other hand, the methods used have adaptations to different behaviors carried out by competitive organizations. To optimize profit, prediction based on user profile is used using robust nonparametric regression.

3. Research and Methodology

In the research methodology, the first step taken is collecting data from mobile wallets based on customers registered in a competitive payment channel. The resulting data is an e-metric data. From the e-metric, the data is grouped based on location, the same merchant type, and different merchant types. This is intended to see customer behavior when transacting at a merchant that results in competition from each merchant. The difference between transactions that occur at merchants is used robustly as an optimization.

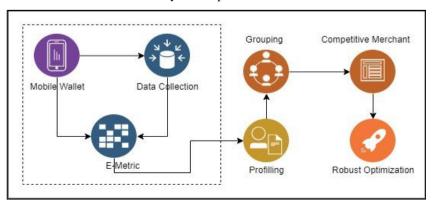


Fig 1 Research methodology

Data collection

This exploration starts by gathering information and data about the quantity of clients and the quantity of vendors in all areas in North Sumatra. The information is caught utilizing a versatile wallet that has been enlisted at different dealers all through North Sumatra [20]. A portable wallet is the utilization of exchange administrations with electronic cash (e-cash) utilizing a wireless or versatile device. The connection among business and purchasers turns out to be quicker utilizing advanced [21] [22]. Associations are changing the way in which they work to meet the ever-changing necessities of their clients and the growing usage of phones and contraptions. Associations in the budgetary territory have hopped up offering mechanized stages and courses of action and are seen as people from the Fintech region [24]. These rising associations are making risky devices and organizations that are viably open at low costs. One of the areas of the budgetary business that is overflowing with improvement is the portion territory.

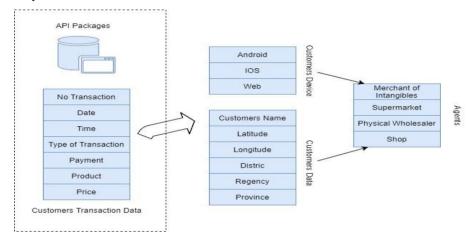


Fig 2 The Mobile Wallet Ecosystem

The important inspiration that drives customer relationships with the system is to combine setup, promotion, and customer service. So, these systems mostly have a dashboard which provides an overall perspective on the ease of transaction.

E-Metric data

Electronic-Metrics or so-called e-metrics are electronic-based customer behavior (e-customer behavior). The technological era has brought big changes to society, this can be seen in presenting everything electronically which provides convenience about how a customer behaves electronically and provides information that we can apply.

m 11	-	T) (
1 ohlo		E-Metric
Lavic	т	L-IVICUIC

Merchant	Freq	Amount
Sate Jo Andah	31	630,000
Toko Baju Anak Anissa	27	203,000
Kemala Sari Kue	28	725,500
Warung Simpang Empat	21	435,000
Bengkel Milala	15	223,500
Irwan Seluler	13	189,000
Warung Jelita	15	265,500
Rujak Danil	17	412,001
Maria Baju	11	154,000
Abu Bakar Ayam Goreng	13	245,000
Febrian Bengkel Las	5	76,000
Sostel Alkudus	7	89,000
Gorengan Maknyus	5	76,500
KFC Haji Adam Malik	8	90,000
Bubur Ayam Jakarta	3	65,500
Burger Malik	4	72,000

4. Profiling Data

With a variety of merchants in utilizing digital business development electronically, it can make it easier to see revenue conditions or what is known as a profitable difference. With a large number and various variations, uncertainty in business is also more difficult to predict. At this stage the merchant variants are grouped and classified so that later there will be outliers or differences that can be used as an advantage at the merchant.

Customer purchases can be represented by several variables which can be seen in table 2 below:

Table 2 Input Variables

INPUT VARIABLES					
X1	Merchant Name				
X2	Customer's name				
X3	Transaction No				
X4	Transaction date				
X5	Product				
X6	Product Prices				

5. Result and Discussion

Grouping

Every object or entity contained in the universe of speech has characteristics on its features and weights. When there is an object whose habit is known, then other habitual behaviour can be efficiently determined based on the measurement of the similarity or closeness between these objects. In topology, x and y are vectors that have magnitude and direction, and A is a vector space, and z is a function of distance for x, y, z U \in A applies z (x, y) = z (x, y) - z (y, i) and x = y. Thus, (X, z) denotes the distance space for the vector space X with the distance function d. The relationship between similarity and inequality as in the equation reveals that they are complementary, that is s (x, y) z = 1 - s (x, y). Referring back to Proposition 1, the points s (x, y) on the straight line L are dual so that if s (x, y) \in [0, 1] then d (x, y) \in [1, 0], where the straight line is L: {s (x, y) I | i = 1, ..., n} \rightarrow [0, 1] while both are straight lines Lc: {d (x, y) I | i = 1, ..., n} \rightarrow [1, 0]. Two lines that are formed will reveal the characteristics of the similarity again as the identity of the similarity itself. After each variable has an appropriate category, a segment is created from the intersection of the valuesx = y or x \cap y> 0.

In table 3, we can see the groupings of merchants who have similarities and habits carried out by customers against competitive merchants.

Merchant P		Gr	oupi	ing														
Sate Jo Andah	2	1	2															
Toko Baju Anak Anissa	1	1																
Kemala Sari Kue	6	1	2	3	4	5	6											
Warung Simpang Empat	8	1	2	3	4	5	6	7	8									
Bengkel Milala	3		2		4	5												
Irwan Seluler	13	1	2	3	4	5	6	7	8	9	10	11		13	14			
Warung Jelita	10	1	2	3	4	5	6	7	8	9		11						
Rujak Danil	10	1	2	3	4	5	6	7	8	9		11						
Maria Baju	16	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
Abu Bakar Ayam	7		2	3	4	5	6	7				11						
Goreng																		
Febrian Bengkel Las	7		2	3	4	5	6	7				11						
Sostel Alkudus	11	1	2	3	4	5	6	7	8	9		11		13				
Gorengan Maknyus	11	1	2	3	4	5	6	7	8	9		11		13				
KFC Haji Adam Malik	17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Bubur Ayam Jakarta	17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Burger Malik	17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

 Table 3 Merchant Groupings

Here each response category is coded via indicator variables. So, if it has a class, then there will be an indicator B_y , Y = 1, ..., Y with $B_y = 1$ if X = Yelse 0. These are pooled together in a vector $B = (B_1, ..., B_k)$, and

this NxY training example forms an indicator response matrix. are matrices of 0 and 1, with each row 1 (1) We adjust the linear regression model to each of the columns from the same time, and the fit is given by

$$B = A(A^T A)^{-1} A^T B$$

Table 4 Behavior based on transaction value

Variabel	Nilai
Transaction	1500; 3000; 3001; 4200; 5500; 6500; 7000; 8000; 9000; 9500;10000; 12000; 13000;
	14500; 15000; 16000; 18500; 19000;20000; 24000; 25000; 27500; 29500; 30000; 31000;
	36000;40000; 42500; 45000; 48000; 50000; 60000; 70000; 72000;75000; 85000; 90000;
	100000; 140000; 185000
Name of Merchant	Sate Jo Andah; Toko Baju Anak Anissa; Kemala sari kue; Warung simpang empat; Bengkel Milala; Irwan Seluler; Warung jelita; rujak danil; Maria Baju; Abu bakar ayam goreng; Febrian bengkel las; Sostel Alkudus; Gorengan Maknyus; KFC Haji adam malik; Bubur ayam Jakarta; Burger malik

6. Robust Optimization

In this study, robust regression of MM (Method of Moment) estimation is used as a method of overcoming outliers on the data used. The results of the identification that have been made in the previous stages show that there are outliers in the data. Furthermore, to overcome this problem, robust regression with MM (Method of Moment) estimation is used. The settlement steps are as follows:

1. Calculating the initial estimator of the coefficients and residuals $\beta_i \varepsilon_i$

2. Calculate the initial regression parameters so that the value is obtained and get the value to find the initial weighted value. $\beta_0 \varepsilon_i = Y_i - Y_i$

3. Look for value
$$\sigma = \sqrt{\frac{n \sum_{i=1}^{n} (\varepsilon_i^2) - (\sum_{i=1}^{n} \varepsilon_i)^2}{n(n-1)}}$$
(2)

4. Calculating value
$$w(u_i) = \frac{\psi(u_i)}{u_i} = \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 |u_i| \le c, \ 0 \ |u_i| > c$$
 (3)

Where $u_i = \frac{\varepsilon_i}{\delta} danc = 1,547$

5. Finding the weighting value as the initial weighting value using the criteria of the Tukey Bisquare weighting function. $w(u_i)$

$$\beta_m = \arg\min\sum_{i=1}^n \rho\left(\frac{y_i - \sum_{j=0}^k x_i \beta_j}{\delta}\right) \tag{4}$$

6. Repeating steps a through d until a convergent value is obtained, namely the difference in value and until it approaches 0.

- 7. Calculating the residual value $\varepsilon_i = Y_i Y_i$
- 8. Calculates the weighted value as the initial weighted value using

$$w(u_i)w(u_i) = \left[1 - \left(\frac{u_i}{c}\right)^2\right]^2 |u_i| \le c, 0 |u_i| > c$$
⁽⁵⁾

and using Tukey Bisquare's weighting function criteria 21^2

$$w(u_i) = \left[1 - \left(\frac{u_i}{c}\right)^2\right]^{-1} |u_i| \le c_i, \text{ Where 0 } |u_i| > cu_i = \frac{\varepsilon_i}{\delta} danc = 4,685$$
(6)

9. Looking for robust regression estimation-MM parameter values

$$\beta = (X^T W^I X)^{-1} X^T W^I y i = 1,2,3 \dots$$

Where:

$$y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_n \end{bmatrix}, X = \begin{bmatrix} 1 & X_1 & X_1 \cdots X_{1k} \\ 1 & X_2 & X_2 \cdots X_{2k} \\ 1 & X_{n1} & X_{n2} \cdots X_n \end{bmatrix}$$
(7)

$$X^{T} = \begin{bmatrix} X_{1} & X_{2} & \dots & X_{n1} \\ X_{1} & X_{2} & \dots & X_{n2} \\ X_{1k} & X_{2k} & \dots & X_{n} \end{bmatrix}$$
$$W = \begin{bmatrix} W_{1,1} & 0 & \dots & 0 \\ 0 & W_{1,2} & \dots & 0 \\ 0 & 0 & \dots & W_{1,n} \end{bmatrix}$$

The iteration will stop until a convergent value is obtained, namely the difference in value and approaches $0\beta_j\beta_j^{i+1}\beta_j^i$

10. Repeating steps 2 to 4 until a convergent value is obtained, namely the difference in value and until it approaches 0.

	Table 5 Robust Parameter Value Iteration Results using a tukey	
me	Difference in value	Value

Merchant Name	Difference in value	Value Comparison
Sate Jo Andah; Toko Baju Anak Anissa	11.5	45.1
Toko Baju Anak Anissa; Kemala Sari Kue	32.8	23.1
Kemala Sari Kue; Toko Baju Anak Anissa	21.8	22.5
Warung Simpang Empat; Kemala Sari Kue	6.6	12.5
Bengkel Milala; Toko Baju Anak Anissa	54.5	11.67
Irwan Selule; Kemala Sari Kue r	78.5	21.6
Warung Jelita; Toko Baju Anak Anissa	44.9	33.6
Rujak Danil; Kemala Sari Kue	21.2	35.2
Maria Baju; Rujak Danil	9.5	47.8
Abu Bakar Ayam Goreng; Rujak Danil	31.3	26.6
Febrian Bengkel Las; Abu Bakar Ayam Goreng	11.3	56.7
Sostel Alkudus; Rujak Danil	45	77.5
Gorengan Maknyus; Abu Bakar Ayam Goreng	34.2	56.6
KFC Haji Adam Malik; Abu Bakar Ayam Goreng	67	12.5
Bubur Ayam Jakarta; Rujak Danil	42	58.5
Burger Malik; Abu Bakar Ayam Goreng	45.1	46.7
Iteration Results	23.9	53.1

Based on Table 3, it can be seen that the difference in estimation between food sales is more pronounced among salesmen who open, repair shops, cell phone shops and grocery stalls. This shows that the parameter estimates have converged.

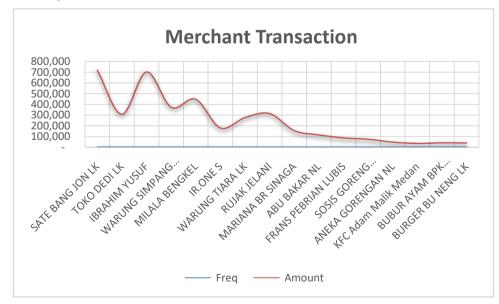


Fig 3 Comparison Results of merchant transaction iterations

Figure 3 shows that the result of iteration of outlier comparisons that occurs in each transaction between competitive merchants is that selling types of food has a higher value.

7. Conclusion

Advancement is influenced by boundary vulnerability, network and the need to deal with unsure information [23]. There are a few wellsprings of vulnerability for example information issues are not known with sureness or may not be estimated unequivocally, or the specific answer for the issue may not be pertinent because of characteristic mistakes of the gadget. Information vulnerability produces dubious imperatives and target capacities [24]. Robust is a displaying procedure for preparing issues where advancement information is unsure and just known to have a place with some arrangement of vulnerabilities. MM (Method of Moment) estimation is one of the estimates in robust regression that can be used for detected outlier data on the independent and dependent variables. Based on the parameter significance test conducted, it was found that merchants who sell food are more profitable than other merchants. With a difference in iteration value 23.9 and a comparison of iteration value 53.1.

References

- El Sawy, OA, & Pereira, F. (2013). Business modelling in the dynamic digital space: An ecosystem approach (Vol. 1, Springer Briefs in digital spaces). Heidelberg: Springer-Verlag
- ST Asah, AD Guerry, DJ Blahna, JJ Lawler, Perception, acquisition and use of ecosystem services: Human behaviour, and ecosystem management and policy implications. Ecosystem Services. 10, 180–186 (2014).
- Business Intelligence and Analytics in Small and Medium Enterprises. Manufacturing Design and Technology Series Series Editor: J. Paulo Davim, University of Aveiro, Portugal, (2020)
- Melo, PN, & Machado, C. (Eds.). (2019). Business Intelligence and Analytics in Small and Medium Enterprises. CRC Press
- Dijkman, Remco M., Marlon Dumas, Boudewijn F. van Dongen, Reina Uba and Jan Mendling. "Similarity of business process models: Metrics and evaluation." Inf. Syst. 36 (2011): 498-516
- Zhang, X., Qin, Z., Liu, X., Hou, Q., Zhang, B., & Wu, J. (2015). Hadoop-Based Similarity Computation System for Composed Documents
- Bardicchia, Marco (2020). Digital CRM: Strategies and Emerging Trends: Building Customer Relationship in the Digital Era. p.12
- C. Tat Huei, L. Suet Cheng, L. Chee Seong, A. Aye Khin, R. Ling Leh Bin, Preliminary Study on Consumer Attitude towards FinTech Products and Services in Malaysia. International Journal of Engineering & Technology. 7, 166 (2018)
- F. Amato et al., In Smart Innovation, Systems and Technologies (Springer Science and Business Media Deutschland GmbH, 2019), vol. 98, pp. 53–63
- R. Syah, MKM Nasution, E. B Nababan, S. Efendi. Knowledge Acceleration Estimator (KAE) Model to Customer Behavior Using Business Metrics. Journal of Theoretical and Applied Information Technology (JATIT 2020)
- ST Asah, AD Guerry, DJ Blahna, JJ Lawler, Perception, acquisition and use of ecosystem services: Human behavior, and ecosystem management and policy implications. Ecosystem Services. 10, 180–186 (2014
- Özmen, A. (2016). Robust Optimization of Spline Models and Complex Regulatory Networks (Vol.139). Switzerland: Springer International Publishing
- Özmen, A., & Weber, GW (2014). RMARS: robustification of multivariate adaptive regression spline under polyhedral uncertainty. Journal of Computational and Applied Mathematics, 259, 914-924
- Özmen, A. (2016). Real-world application with our robust tools. In Robust Optimization of Spline Models and Complex Regulatory Networks (pp. 89-113). Springer, Cham
- Syah, R., Elveny, M., & Nasution, M. K. (2020, October). Clustering Large DataSet'to Prediction Business Metrics. In Proceedings of the Computational Methods in Systems and Software (pp. 1117-1127). Springer, Cham.
- Özmen, A., Weber, GW, Batmaz, İ., & Kropat, E. (2011). RCMARS: Robustification of CMARS with different scenarios under polyhedral uncertainty set. Communications in Nonlinear Science and Numerical Simulation, 16 (12), 4780-4787M. KM Nasution, OS Sitompul, S. Nasution, H. Ambarita, New Similarity, IOP Conf. Series: Materials Science and Engineering 180 (2017) 012297
- Özmen, A., Weber, GW, & Batmaz, İ. (2010). The new robust CMARS (RCMARS) method. vectors, 1, 2
- Susanti, Y., & Pratiwi, H. (2014). M estimation, S estimation, and MM estimation in robust regression. International Journal of Pure and Applied Mathematics, 91 (3), 349-360.
- M Elveny., R Syah., MKM, Nasution. 2020. Taxonomy Business Model to Customer Behavior in E-Metric Ecosystem. Journal of Advanced Research in Dynamical and Control Systems. 12 (7), pp. 99-104. DOI: 10.5373 / JARDCS / V12I7 / 20201989
- GW Weber, B Kjamili, D. Czerkawski, 2019. LiBerated Social Entrepreneur. Using Business Metrics: Migport Refugee Big Data Analytics. With a Note on Ability and Disability. Engineering Science and Technology, an International Journal, JESTECH (Elsevier Inc, .2019).
- El Ghaoui, L., Lebret, H. (1997): Robust solutions to least-square problems with uncertain data matrices. SIAM J. Matrix Anal. Appl. 18, 1035–1064

- C. Tat Huei, L. Suet Cheng, L. Chee Seong, A. Aye Khin, R. Ling Leh Bin, Preliminary Study on Consumer Attitude towards FinTech Products and Services in Malaysia.International Journal of Engineering & Technology. 7, 166 (2018).
- Syah, R., Elveny, M., & Nasution, M. K. (2020, June). Performance Knowledge Acceleration Optimization with MARS to Customer Behavior in Merchant Ecosystem. In 2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT) (pp. 178-182). IEEE.
- Ben-Tal, A., Margalit, T., Nemirovski, A. (2000): Robust modeling of multi-stage portfolio problems. In: Frenk, H., Roos, C., Terlaky, T., Zhang, S., eds., High Performance Optimization. Kluwer Academic Publishers, pp. 303–328
- Ben-Tal, A., Nemirovski, A., Roos, C.: Robust solutions of uncertain quadratic and conic quadratic problems. Research Report # 2/01, Minerva Optimization Center, Technion - Israel Institute of Technology, Technion City, Haifa 32000, Israel.http://iew3.technion.ac.il:8080/subhome.phtml?/Home/research
- A. Goli, HK Zare, R. Tavakkoli-Moghaddam and A. Sadeghieh, Application of robust optimization for a product portfolio problem using an invasive weed optimization algorithm, Numerical Algebra, Control & Optimization, 9 (2019), 187–209.
- Syah, R., Elveny, M., Nasution, M. K., & Weber, G. M. (2020). Enhanced Knowledge Acceleration Estimator Optimally with MARS to Business Metrics in Merchant Ecosystem. In 2020 4rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM) (pp. 1-6). IEEE