

Electronic Commerce Product Recommendation using Enhanced Conjoint Analysis

Andrew Brian Osmond, Fadhil Hidayat, Suhono Harso Supangkat

School of Electrical Engineering and Informatics

Center of Smart City Community and Innovation, Institut Teknologi Bandung, Indonesia

Abstract—While finding any product, there are many identical products sold in the marketplace, so buyers usually compare the items according to the desired preferences, for example, price, seller reputation, product reviews, and shipping cost. From each preference, buyers count subjectively to make a final decision on which product is should be bought. With hundreds of thousands of products to be compared, the buyer may not get the product that meets his preferences. To that end, we proposed the Enhanced Conjoint Analysis method. Conjoint Analysis is a common method to draw marketing strategy from a product or analyze important factors of a product. From its feature, this method also can be used to analyze important factors from a product in the marketplace based on price. We convert importance factor percentage as a coefficient to calculate weight from every attributes and summarize it. To evaluate this method, we compared the ECA method to another prediction algorithm: generalized linear model (GLM), decision tree (DT), random forest (RF), gradient boosted trees (GBT), and support vector machine (SVM). Our experimental results, ECA running time is 6.146s, GLM (5.537s), DT (1s), RF (10,119s), GBT (45.881s), and SVM (11.583s). With this result, our proposed method can be used to create recommendations besides the neural network or machine learning approach.

Keywords—Enhanced conjoint analysis; marketplace; e-commerce

I. INTRODUCTION

e-Commerce, in its first form, was established 40 years ago. Since then, new technology, advances in internet connectivity and security, payment gateways, and widespread consumer and business model acceptance have all aided e-growth. Commerce's Michael Aldrich launched electronic shopping in 1979 by connecting a customized television to a transaction-processing computer through a telephone line. Amazon was one of the earliest e-commerce sites for books, and PayPal followed it as an e-commerce payment system in 1995. Alibaba was founded in 1999 as an online e-commerce platform. It transforms the company model from single-store e-commerce to a multistore marketplace.

In the 2000s, Shopify and BigCommerce have launched a cloud e-commerce platform that helps retailers to have their e-commerce shop. Previously our system had been built from scratch, but Shopify and Bigcommerce allow us to have our e-commerce store by renting their platform. Since 2010, numerous payment systems have arisen, including stripe, apple pay, Samsung Pay, and other payment method. Now, there is a social commerce, where people can buy or sell something from

social media like Facebook, Instagram, and even messenger applications like WhatsApp, telegram, and signal.

In Indonesia, Bhinneka and Sanur are the first ecommerce in Indonesia in 1996. Then Doku as payment gateway system is introduced in 2007. New era of ecommerce starts from 2010 as Tokopedia and Bukalapak launch as e-commerce marketplace. Later, there are a lot of marketplace launched such as: Blibli, JD.ID, Lazada, qoo10, and many more. e-Commerce is a new way in business transaction through internet with cover of lease or the auction goods [1]. e-Commerce marketplace in Indonesia are commonly act as mediator. Marketplace platform act as bridge between customer and seller, they obtained data generated by all its participant [2]. Author [3] formulate e-commerce as integrated platform from upstream supplier (individual industry, cooperative, and production) into end customer. This platform consists of intermediate channel provider (distributor, importer), one or few categories of product, called vertical e-commerce, and O2O (online-to-offline) model where customer can browse the product through e-commerce and buy it in physical store.

e-Commerce today is not only a rigid platform for interaction between buyers and sellers, where information received by buyers only comes from the platform but provides direct access between buyers and sellers to create an interactive environment [4]. Even e-commerce offers transactions in one country and more than one country; it is called cross border e-commerce, a new concept about resource integration, supply and demand matching, joint operation, and supply chain logistics between countries [5]. People buy some items at online stores for a variety of reasons: They can buy at any time. They save time by not having to go to the store. They can easily compare prices. They do not have to bargain because the price of the product is fixed. More information and chances to compare goods and prices among thousands of products are available to customers in an online shop, and better product selection, convenience, and simplicity of discovering desired products [6]. In the marketplace, consumers usually search the products and filter them according to the needs of consumers, for example, location, price, discount, delivery courier. All sellers with their products will appear based on appropriate keywords. After choosing the product, consumers will choose the delivery courier and payment method. Finally, after payment is completed, consumers can wait for their product to come. Consumers can finish their order when the product comes, and the fund can be released to the seller. Flow of e-commerce transaction is shown in Fig. 1.

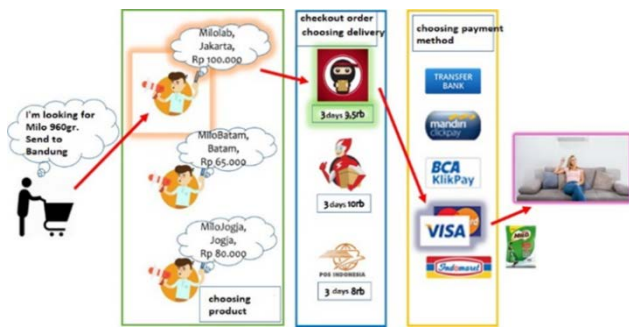


Fig. 1. Flow of e-commerce Transaction.

Currently, transactions in a marketplace are determined by many factors:

- 1) Location of buyers and sellers. The location of the buyer and seller usually becomes one of the things that are considered for buying goods because relating to shipping cost.
- 2) Delivery services that the seller activates. Every buyer has an opinion or experience using certain shipping services.
- 3) Price. One of the essential aspects for purchasers is the pricing of goods. Buyers hunt for the best deal on everyday items.
- 4) Review. Reviews about products, services, and shipping from certain transactions of previous buyers.
- 5) Promotion. Promotions are provided by sellers, such as discounts, cashback, or free shipping.

Indonesia is regarded as one of Asia's most developed e-commerce markets. Several reasons contribute to Indonesia's e-commerce sector's fast growth. To begin with, smartphone and Internet development continue to accelerate. Second, Indonesia has a large population with rising purchasing power due to the country's macroeconomic solid growth. Third, Indonesia's populace is young and tech-savvy. [2]. There are 175.4 million active Internet users in January 2020 [3]. With 2.201 start-ups, Indonesia can adjust and develop product based new technology rapidly [4]. Besides, the support from the government by opening the large scale of investment from domestic or overseas to participate in developing Micro, Small and Medium Enterprises and e-commerce have made e-commerce grow very fast. This growth gives chances to e-commerce practitioners to grow their business, such as Tokopedia, Bukalapak, Blibli, and others [5]. In Indonesia, buyers usually spend an average of 4 minutes in a marketplace to have a look at some products before deciding to buy an item or not. In general, there are many identical products sold in the marketplace, so buyers usually compare the items according to the desired preferences, for example: price, seller reputation, product reviews, and shipping cost. From each preference, buyers count subjectively to make a final decision which product is should be bought. With so many products in e-commerce, there is a problem finding the optimal product because buyers have to manually compare products one by one.

Motivated by the problem, the following research the question arises and will be discussed in this paper: what types of consumer strategy should be implemented to get the optimal product?

To answer the question above, we proposed a method for helping buyer finds optimal product from price, shipping cost, and insurance cost. An enhanced conjoint analysis method is chosen to recommend optimally product. Conjoint analysis is a well-known research technique in marketing and consumer research. This the approach has been used to address a wide range of marketing challenges, including predicting product demand, developing a new product line, and calibrating pricing sensitivity/elasticity. The approach entails presenting respondent consumers with a carefully crafted collection of hypothetical product profiles (defined by the necessary levels of relevant qualities) and gathering their preferences for those profiles in the form of ratings, rankings, or selections [6].

Conjoint analysis usually is used to identify some variables that important for a product. For example: [7] using adaptive choice-based conjoint analysis to identify surcharge for outdoor apparel, [8] evaluate domestic express coach service using conjoint analysis, [9] identify the critical attribute of smartphones using conjoint 3 analyses. Widely conjoint analysis is used to evaluate essential attributes for a product. By knowing the critical attribute of the product, a company can create a new marketing strategy, new version of the product, or new product that is close to what consumer needs. In this research, a conjoint analysis is proposed to get essential attributes from a consumer's product. Value from attribute is converted as attribute weight. Later multiply the value of each product attribute and its weight to get a score and choose the minimum score as a recommendation. Assumed that there exists e-commerce. When the consumer chooses the product, our system can recommend what the consumer should choose.

The following is a breakdown of the paper's structure. The second section examines research in a similar field. We introduce notations, assumptions, formulations, and the system that will be used to answer the problems in Section 3. To identify the performance of our method, we create simulation, test, and analyze it. Section 4 contains the conclusion and future works of this research.

II. LITERATURE REVIEW

Identify applicable sponsor/s here. If no sponsors, delete this text box (sponsors). Some study has been done on pricing strategies in the marketplace: [10] offers an empirical analysis and analytical a model that shows how an online shop may achieve a competitive advantage by designing an optimal mix of product price and shipping price. In [11], has proposed a unique algorithmic approach for estimating optimal prices in e-commerce scenarios using noisy and sparse data. Their structure is broad and can be tailored to a particular issue. Experiments have demonstrated that using their methodology in practice can result in significant increases in profit and revenue. In [12], recommends a multi-armed bandits algorithm for dynamic pricing using a customer-centric strategy, based on the notion that systems track client behavior and price impacts whether or not a purchase is made. In the e-commerce recommender system, there is some research about it. Collaborative filtering, knowledge-based reasoning, content-based filtering, demographic-based filtering and hybrid technique are still become a favorite primary method to be explored. In [13], improve recommender system using previous

search information, user behavior analysis, and current search information. The proposed method by combining content and web usage mining technique then measure the accuracy of the system. In [14], improving collaborative filtering to recommend trendy items to a customer to demonstrate the performance of their algorithm, they took the system to an actual retail mall. They claim that their system outperforms traditional collaborative filtering in terms of efficiency. In [15], developed and implemented a recommender system based on web mining. Their complicated recommendation engine takes data from web mining as input. For the implementation of a recommender system, [16] employed demographic data from the client registration form as an essential source of data mining algorithm. They proposed four-phase in their recommender system workflow: 1) acquire information implicitly and explicitly, 2) information processing, 3) recommendation processing, Moreover, 4) indicating the outcome to the clients. They also employed a data diversity such as online market data, query data, server log data, hyperlink inside web pages, client registration data, and web pages. From the research, the author concludes that the proposed approach improves the quality of recommendation efficiently. To construct their recommendation system, [17], [18], [19] used a knowledge-based filtering strategy. In [19], Using KBF to help a customer make a decision. They used the length of time spent shopping and the kind of things purchased by the consumer as input to choose the most important product. It was discovered that KBF could improve decision-making by reducing the length of shopping and the effort required to choose the right product. In [18], using data clustering analysis results to develop a web mining framework. It is used to deliver recommendations for e-commerce recommender systems. He also stores knowledge rules using pattern analysis on acceptable media and learns the recommender engine from web mining. The author concluded that the suggested method could manage massive data quantities while increasing reaction time substantially and scalability. There is also some research about recommender systems in e-commerce using content-based filtering methods. In [20], suggest content analysis for improving the recommender system. The proposed method will recommend sample websites to meet user requirements. However, it has a high operation cost and response time because many websites have different structures, naming tags, and information. In [21], introduced BPR (Bayesian Personalized Ranking) to explore user preferences by ranking. They employed a global score function to determine the user's preference for various goods. In [22], proposed a strategy that used CF and CBF to achieve excellent recommendation accuracy. The author proposes a three-part weighted combination filtering recommender system: gathering required data, creating recommendations, and communicating findings to the user. Moreover, one of the most popular recommender system research techniques is a hybrid technique by joining various algorithms and methods. Several different algorithms are used for collecting and processing information that improves the quality of recommendations. In [23], developed a weighted parallel technique for developing an e-commerce recommender system in Indonesia. The author uses a combination of CF and CBF approaches. While in [24] uses a personalized hybrid way to evaluate standard algorithms to

enhance user interest modelling, variety, and scale. As we know so far, there is no research about using the Conjoint Analysis method to create a recommender system, especially in e-commerce. So, we assess CA as a base method to create recommendations, especially for e-commerce in Indonesia.

III. NOTATION, FORMULATION, AND SYSTEM DESIGN

For convenience, we summarize the notation adopted in this paper is in Table II. In this research, we do preliminary study to determine what kind of attribute to use in this research. We investigate three of e-commerce in Indonesia: Tokopedia (<https://www.tokopedia.com>), bukalapak (<https://www.bukalapak.com>), and shopee (<https://shopee.co.id>). Each explicit product attribute is shown in Table I.

TABLE I. PRODUCT ATTRIBUTES

	Tokopedia	Bukalapak	Shopee
Product Name	✓	✓	✓
Sold	✓	✓	✓
Weight	✓	✓	✓
Tags		✓	
Percentage of satisfaction consumer		✓	
Order processing time		✓	
Insurance Cost		✓	
Price	✓	✓	✓
Free shipping badge	✓		✓
Delivery courier	✓	✓	✓
Buyer Location	✓	✓	✓
Stock	✓	✓	✓
Description	✓	✓	✓
Categories	✓	✓	✓
Brand	✓	✓	✓
Seller Location	✓	✓	✓
Review score	✓	✓	✓

In addition to determining the essential qualities, the examinable performance level for each attribute must be defined using the following [25] criteria: 1) attribute levels should be as similar to the real-life experience of the customer as feasible, 2) attribute levels should be related to the product that is accessible to the consumer, and 3) contain characteristics that are regarded to be essential competences. Then determine that the primary attributes that will be used for this method should be numeric.

Product score, sold, price, shipping cost, and insurance cost are chosen as our method's primary attributes. When not all preferences are satisfied, the customer must make a trade-off between all primary attributes during the selection process. We try to solve it using Multi Attributes Decision Making (MADM) problem formulation. The problem formulation for MADM is as follows [26]:

- 1) There is a set of X attributes, where $X = \{X_1, X_2, \dots, X_n\}$
- 2) a set of domain values $D = \{D_1, D_2, \dots, D_n\}$: each $D_i (1 \leq i \leq n)$ represents a collection of possible values for attribute X_i . To prevent confusion, we use D to represent the space of all possible outcomes.

3) a set of restrictions $C = \{C_1, C_2, \dots, C_m\}$: where each $C_j(1 \leq j \leq m)$ is a constraint function that limits the possible values for a subset of the characteristics X.

4) a list of available outcomes $O = \{O_1, O_2, \dots, O_l\}$: where each $O_j(1 \leq j \leq l)$ is an element of the possible result space D, and O is a subset of D. Although the size of O is usually smaller than that of D, it is still too large for the decision maker to choose one at a time. This set must contain the solution(s) that the decision maker ultimately chooses.

5) a set of statements expressing the decision maker's preferences $P = \{P_1, P_2, \dots, P_t\}$: this piece of data must be elicited from the decision maker prior to or during the encounter. Different decision makers may have different preference statements, and during the process of searching for the best answer, certain preferences may be violated for tradeoff purposes.

In this paper, consider there are i product ($i = 1, 2, 3, \dots, N$). Every product i has their attributes product score (c_i), sold (d_i), price (p_i), shipping cost (s_i), and insurance cost (u_i). Each of attributes has a weight coefficient denoted by K, L, M, N, O . Notation is shown in Table II.

Conjoint analysis is a well-known research technique in marketing and consumer research. This methodology, which allows for the understanding of consumer preferences, has been used to solve a wide range of marketing challenges, including forecasting product demand, creating a new product line, and calibrating pricing sensitivity/elasticity. Respondent customers are presented with a well-crafted collection of hypothetical product profiles (defined by the stated levels of the necessary qualities), and their preferences are collected in the form of ratings, rankings, or selections for those profiles [6]. The technique of ordinary least square regression (OLS) is often used to estimate preference functions. According to research, this technique's efficiency (predictive power) is often relatively similar to more complex techniques, but the findings are easier to grasp. OLS equation is shown below:

$$S = \sum_{i=1}^n (y_i - \tilde{y}_i)^2 = \sum_{i=1}^n (y_i - b_1x_1 - b_0)^2 = \sum_{i=1}^n (E_i)^2 = \min$$

where,

\tilde{y}_i = predicted value for i observation,

y_i = actual value for i observation,

E_i = error / residual for i observation,

n = total number of observation.

Global utility of a given attribute is determined using the equation (2), where Op is the relative significance of the product attribute, $max up$ is the utility of the attribute's most preferred level. $Min up$ is the utility of the attribute's least chosen performance level. The operation is continually changing and based on a variety of factors.

$$Op = \frac{(max(u_p) - min(u_p))}{\sum_{p=1}^t (max(u_p) - min(u_p))} \quad (2)$$

TABLE II. NOTATION

Notation	Description
c_i	Review score for product i
d_i	Number of sold for product i
p_i	Price for product i
s_i	Shipping cost for product i
u_i	Insurance cost for product i
K	Weight for product score
L	Weight for number of sold
M	Weight for price
N	Weight for shipping cost
O	Weight for insurance cost

Here, define that

$$O_1 = K,$$

$$O_2 = L,$$

$$O_3 = N,$$

$$O_4 = O,$$

Then we construct an enhanced equation to calculate weight for all attributes from each product below (3) and find the minimum value for the weight of the i product as an enhanced conjoint analysis method to give a recommendation.

$$R = \min(\sum_{i=1}^n (Kc_i + Ld_i + Ns_i + Ou_i)) \quad (3)$$

A. Designed System

The designed recommender system is shown in Fig. 2. The flow of the recommendation process is described as follows:

1) User the product name as a keyword, then aggregator will collect the data from the marketplace and save it to the database.

2) Pre-processor will get the data from the database and clean it first (such as: removing NULL values, unused attributes) before it comes to the conjoint analyzer.

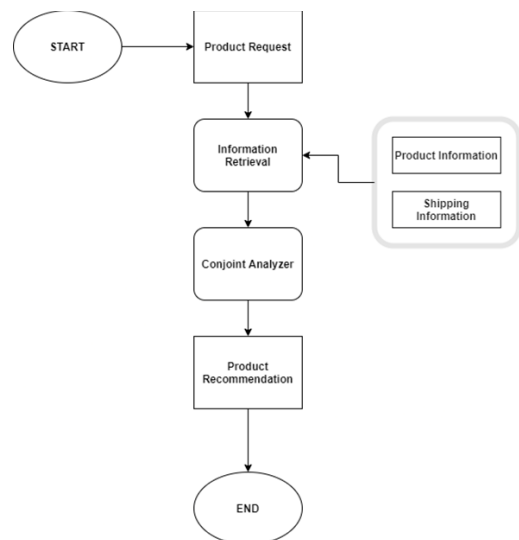


Fig. 2. Designed System.

3) In Conjoint Analyzer, equation (1) is implemented for each attribute ($c_i, d_i, p_i, s_i,$ and u_i) from product i . The result from this function is O of each attributes p .

4) Choosing the minimum value using equation (3) and recommend it to the user.

Table III shows how construct a table to store data from the marketplace.

TABLE III. PRODUCT TABLE

Field	Type	Length
id	Integer	11, PK
Link	Varchar	300
ProductName	Varchar	255
Price	Integer	11
Sold	Integer	11
Reviews	Integer	11
Location	Varchar	30
Delivery Services	Varchar	40
ShippingCost	Integer	11
InsuranceCost	Integer	11
Eta	Varchar	11
etalInfo	Varchar	11

B. Dataset Processing

There are some choices of Indonesian marketplace as data source, such as: Tokopedia, Shopee, Bukalapak, Lazada, Jakmall, or Blibli. Bukalapak has been chosen as our data source because Bukalapak provides API to get information about product delivery instead of product from their website. Assumed that users search for products using “gegep tekiro” to grab products related to “gegep tekiro.” A searching page for the keyword “gegep tekiro” is shown in Fig. 4. Later from the searching page, go to the detail of every product. Information on the detailed product is shown in Fig. 5. We retrieve product name, link, price, sold, reviews, location, delivery services, shipping cost, insurance cost, eta (estimated time arrival), and eta on the detailed product page info.

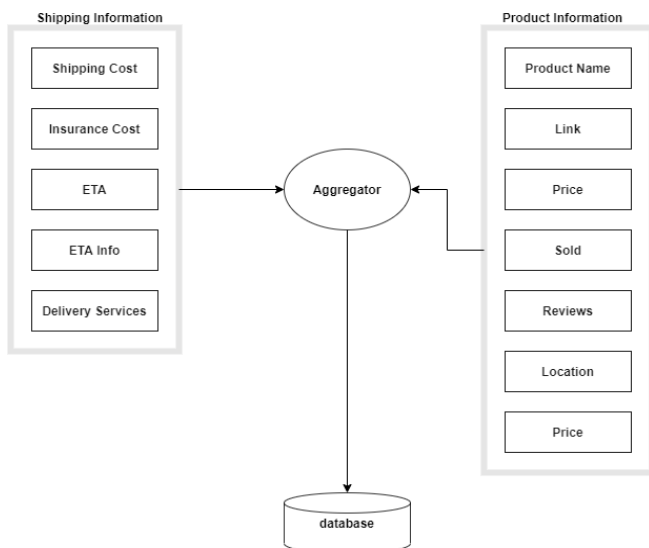


Fig. 3. Parameter Processing.

An aggregator is created to retrieve only needed information based on the HTML tag on the product detail page. Eliminating unessential parameters, then insert the value into the table on the database. Parameters processing is shown in Fig. 3.

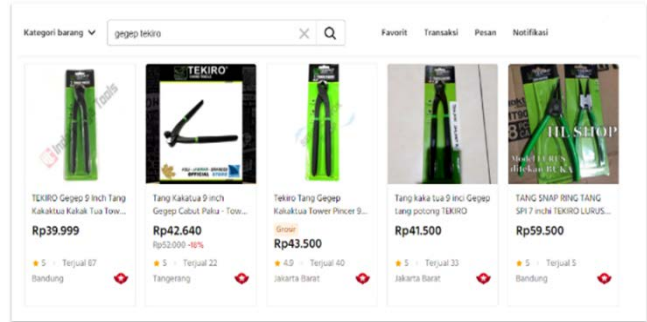


Fig. 4. Product Result Page.

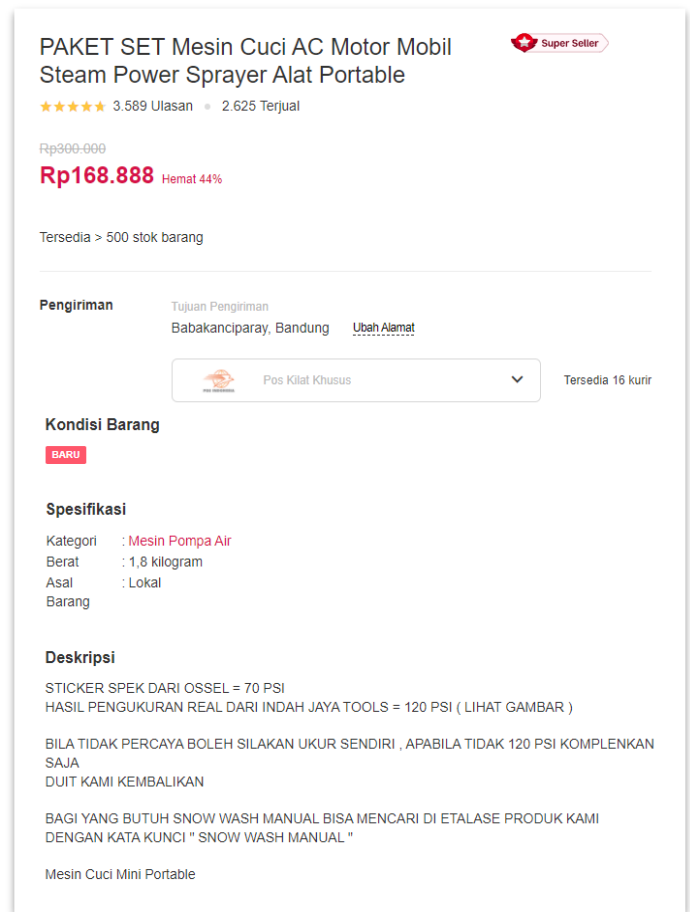


Fig. 5. Product Detail Page.

IV. SIMULATION

In this experiment, we limit data to 240 sellers. The five cheapest product is selected from the dataset. In this simulation, we will simulate the Enhanced Conjoint Analysis method compared to price. There are three scenarios with dynamic growth data starting from 80 data, 160 data, and 240 data. The five cheapest products are shown in Table IV.

TABLE IV. FIVE CHEAPEST PRODUCT BY PRICE

Name	Price (Rp)
tang gegep 9 inch kakatua tower pincer Tekiro	39.300
TEKIRO tang kakatua 9 tang gegep 9 tower pincer 9 inch	39.700
TEKIRO Gegep 9 Inch Tang Kakatua Kakak Tua Tower Pincer	39.999
TANG GESEP - KAKAKTUA 9 INCH TEKIRO - TOWER PINCER	40.000
Tang Kakatua Tang Gegep Tower Pincer 9 TEKIRO	41.000

Three different scenarios are implemented to get result from our method. The result is described as follows:

1) Scenario 1: Importance attribute from conjoint analysis process is shown in Table V. Here, the most important attribute is insurance cost, followed by shipping cost, sold, and reviews.

Keywords	Gegep tekiro
Total seller	80 seller
Location buyer	Bandung

TABLE V. IMPORTANCE ATTRIBUTES OF SCENARIO 1

Attributes	Importance Percentage
Insurance cost	63.1 %
Shipping cost	14.7 %
Sold	12.8 %
Reviews	9.3 %

Later, equation to create recommendation is:

$$R = \min(\sum_{i=1}^n (63.1c_i + 12.8d_i + 14.7s_i + 63.1u_i)) \quad (4)$$

TABLE VI. RESULT OF SCENARIO 1

Name	P	S	R	SC	IC	TW
TEKIRO Gegep 9 Inch ...	39.999	52	5.0	8000	216	1319.42
Tang Gegep Tekiro 9 inches ...	45.000	0	0	8000	238	1326.18
TEKIRO Gegep 9 Inch ...	39.999	52	5.0	8500	218	1394.18
Tang Gegep Tekiro 9 inches ...	45.000	0	0	8500	241	1401.57
Gegep Tekiro 9 in	45.000	12	5.0	8500	241	1403.57

Note: P: price, S: sold, SC: shipping cost, IC: insurance cost, TW: total weight.

In this scenario, the obtained result is different from the product in Table VI. Our system chooses a product with a price of Rp 39.999 with the combination of shipping cost Rp 8.000, insurance cost Rp 216, 52 sold, and have 5.0 reputation as show in Table VI.

2) Scenario 2: Importance attribute from conjoint analysis process is shown in Table VII. Here, the most important attribute is insurance cost, followed by shipping cost, sold, and reviews.

Keywords	Gegep tekiro
Total seller	160 seller
Location buyer	Bandung

TABLE VII. IMPORTANCE ATTRIBUTES OF SCENARIO 2

Attributes	Importance Percentage
Insurance cost	73.8 %
Shipping cost	13.1 %
Sold	8.5 %
Reviews	4.7 %

equation to create recommendation is:

$$R = \min(\sum_{i=1}^n (4.7c_i + 8.5d_i + 13.1s_i + 73.8u_i)) \quad (5)$$

TABLE VIII. RESULT OF SCENARIO 2

Name	P	S	R	SC	IC	TW
TEKIRO Gegep 9 Inch ...	39.999	52	5.0	8000	216	1212.06
Tang Gegep Tekiro 9 inches ...	45.000	0	0	8000	238	1223.64
TEKIRO Gegep 9 Inch ...	39.999	52	5.0	8500	218	1279.04
Tang Gegep Tekiro 9 inches ...	45.000	0	0	8500	241	1291.36
Gegep Tekiro 9 in	45.000	12	5.0	8500	241	1292.61

In this scenario, using twice data more than scenario one, the obtained result is the same as scenario one, but the equation to create a recommendation is different. Numbers of data makes different results for importance factor value and affected variable, used in (5). Our system chooses a product with a price of Rp 39.999 with the combination of shipping cost Rp 8.000, insurance cost Rp 216, 52 sold, and have a 5.0 reputation as show in Table VIII.

3) Scenario 3: Importance attribute from conjoint analysis process is shown in Table IX. Here, the most important attribute is insurance cost, followed by shipping cost, sold, and reviews.

Keywords	Gegep tekiro
Total seller	240 seller
Location buyer	Bandung

TABLE IX. IMPORTANCE ATTRIBUTES OF SCENARIO 3

Attributes	Importance Percentage
Insurance cost	73.4 %
Shipping cost	12.9 %
Sold	8.8 %
Reviews	5.0 %

Equation to create recommendation is:

$$R = \min(\sum_{i=1}^n (5.0c_i + 8.8d_i + 12.9s_i + 73.4u_i)) \quad (6)$$

In this scenario, using three times more data than in scenario 2, our systems still choose products with combinations the same as indicated by scenario 2 as show in Table X. For each scenario, our system will calculate the coefficient of every preference depending on the number of data, and the result can be different. The most interesting is that insurance cost is the most critical factor in a product besides price.

TABLE X. RESULT OF SCENARIO 3

Name	P	S	R	SC	IC	TW
TEKIRO Gegep 9 Inch ...	39.999	52	5.0	8000	216	1195.15
Tang Gegep Tekiro 9 inches ..	45.000	0	0	8000	238	1206.09
TEKIRO Gegep 9 Inch ...	39.999	52	5.0	8500	218	1261.11
Tang Gegep Tekiro 9 inches ..	45.000	0	0	8500	241	1273.39
Gegep Tekiro 9 in	45.000	12	5.0	8500	241	1274.48

For choosing a product in the marketplace, we need more than one attributes to determine which one is optimal in price. We also compared it with another method to observe its performance, relativity error, correlation, and root mean square error. Another method we try to compare is the Generalized Linear Model (GLM), Decision Tree (DT), Random Forest (RF), Gradient Boosted Tree (GBT), and Support Vector Machine (SVM).

From Fig. 6, our method outperforms another method in relative error results. Our experimental results, ECA relative error is 0.0001, GLM (0.008), DT (0.018), RF (0.048), GBT (0.014), and SVM (0.031). Compared to another method, ECA does not need training and test data. Using OLS regression, we obtain an equation and then apply it to the dataset, while another method still needs to predict using training and test data. A relative error has a relation with correlation. With the smallest relative error, the correlation of the ECA method is the highest as show in Fig. 7. ECA correlation is 0.998, GLM (0.994), DT (0.987), RF (0.975), GBT (0.992), and SVM (0.979).

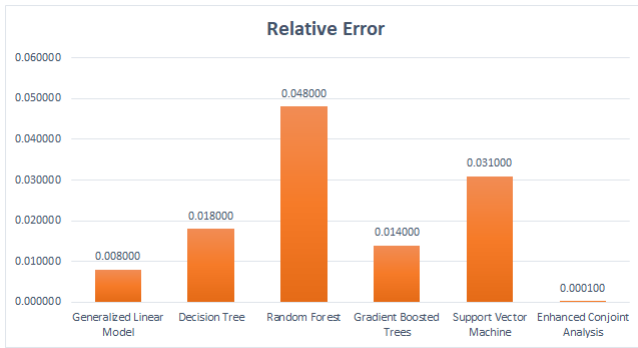


Fig. 6. Relative Error.

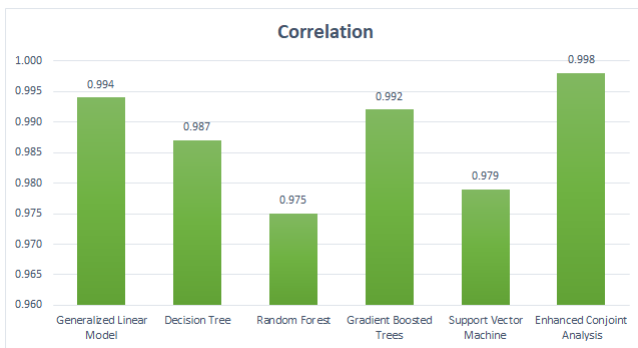


Fig. 7. Correlation.

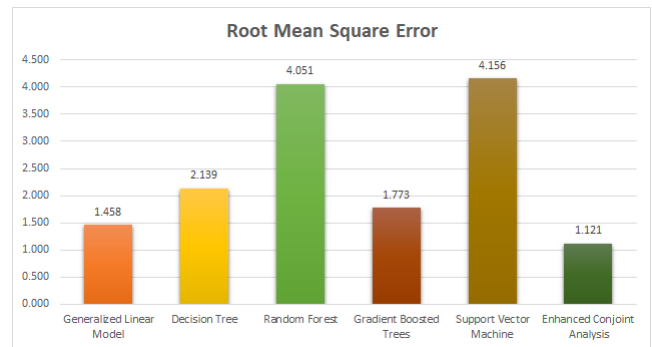


Fig. 8. Root Mean Square Error.

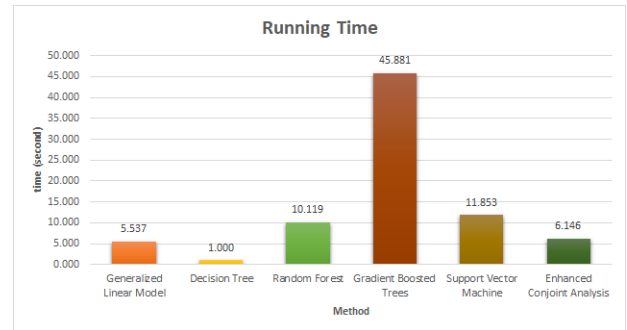


Fig. 9. Running Time.

Since ECA has the slightest relative error, it also has the smallest value for RMSE as show in Fig. 8. ECA RMSE value is 1.121, GLM (1.458), DT (2.139), RF (4.051), GBT (1.773), and SVM (4.156). Although the ECA method has a small error, the decision tree has the best running time as show in Fig. 9. ECA running time is 6.146s, GLM (5.537s), DT (1s), RF (10.119s), GBT (45.881s), and SVM (11.583s). The ECA method has a small difference value from the GLM method because both ECA and GLM are extended linear models. ECA method use Ordinary Least Square as a linear model and GLM using linear regression.

Since we do not have many features in this dataset, the Decision Tree does not have to create a big tree so it can do the deep search faster into the branch, while our method, though it has a regression equation, we still need to calculate a value for each data then sort it. The larger the data, the more time it is needed to calculate.

V. LIMITATIONS AND CONCLUSION

Conjoint Analysis is a standard method to draw marketing strategy from a product or analyze essential factors of a product. This method can also analyze critical factors from a product in the marketplace based on its price. We convert importance factor percentage as a coefficient to calculate weight from every attribute. The simulation results show how that conjoint analysis network works well in choosing a product from multiple attributes to get the best price.

Although our method has good performance, there are some limitations to this method. ECA needs time to calculate a value for each data then sort it. If the data becomes more significant than before, our method needs more time because ECA will calculate all data. We can reduce computing time by

sampling some data to be calculated, but it also can reduce the accuracy result. Another approach is by using a parallel computing approach in calculating value, but it needs observations.

Since ECA uses OLS regression to create an equation, it will choose the optimal value for each attribute. Optimal value can be the biggest or the smallest value in an attribute. In a real-world implementation, if one online shop has a lot of sales, reviews, and low prices, it will be recommended continuously by our method.

Besides its limitation, ECA gives optimal recommendation because all attributes processed using OLS regression give optimal coefficient for every attribute. ECA has the slightest relative error (0.0001) and good correlation and RMSE. The accuracy and the optimization are traded off with computing time. ECA has a running time of 6.146s meanwhile decision tree method has the best time in 1s. With this result, ECA can be used as one alternative recommendation system besides a neural network or machine learning approach.

REFERENCES

- [1] P. Butler and J. Peppard, "Consumer purchasing on the Internet: Processes and prospects," *Eur. Manag. J.*, vol. 16, no. 5, pp. 600–610, 1998.
- [2] A. V. Gunawan, L. Linawati, D. Pranandito, and R. Kartono, "The Determinant Factors of E-Commerce Purchase Decision in Jakarta and Tangerang," *Binus Bus. Rev.*, vol. 10, no. 1, pp. 21–29, 2019.
- [3] S. Kemp, "DIGITAL 2020: INDONESIA," 2020. [Online]. Available: <https://datareportal.com/reports/digital-2020-indonesia>.
- [4] Startup ranking, "Startups per Country," 2020. [Online]. Available: <https://www.startupranking.com/countries>.
- [5] H. Harsono, "Indonesia will be Asia's next biggest e-commerce market," 2017. [Online]. Available: <https://techcrunch.com/2016/07/29/indonesia-will-be-asias-next-biggest-e-commerce-market/>.
- [6] J. Agarwal, W. S. DeSarbo, N. K. Malhotra, and V. R. Rao, "An Interdisciplinary Review of Research in Conjoint Analysis: Recent Developments and Directions for Future Research," *Cust. Needs Solut.*, vol. 2, no. 1, pp. 19–40, 2015.
- [7] B. M. Brand and T. M. Rausch, "Examining sustainability surcharges for outdoor apparel using Adaptive Choice-Based Conjoint analysis," *J. Clean. Prod.*, no. January 2021, p. 125654, 2020.
- [8] M. Walter, F. Haunerland, and R. Moll, "Heavily regulated, but promising prospects: Entry in the German Express Coach Market," *Transp. Policy*, vol. 18, no. 2, pp. 373–381, 2011.
- [9] A. Anand, G. Bansal, and D. Aggrawal, "Choice based diffusion model for predicting sales of mobile phones using conjoint analysis," *J. High Technol. Manag. Res.*, vol. 29, no. 2, pp. 216–226, 2018.
- [10] Y. Yao and J. Zhang, "Pricing for shipping services of online retailers: Analytical and empirical approaches," *Decis. Support Syst.*, vol. 53, no. 2, pp. 368–380, 2012.
- [11] J. Bauer and D. Jannach, "Optimal pricing in e-commerce based on sparse and noisy data," *Decis. Support Syst.*, vol. 106, pp. 53–63, 2018.
- [12] M. Castronovo et al., "Learning Exploration/Exploitation Strategies for Single Trajectory Reinforcement Learning," *Conf. Proc.*, vol. 24, no. 2002, pp. 1–9, 2012.
- [13] V. Bajpai and Y. Yadav, "An Improved Dynamic E-Commerce Recommendation System," *SSRN Electron. J.*, 2019.
- [14] H. Hwangbo, Y. S. Kim, and K. J. Cha, "Recommendation system development for fashion retail e-commerce," *Electron. Commer. Res. Appl.*, vol. 28, pp. 94–101, 2018.
- [15] Y. K. Choi and S. K. Kim, "An auxiliary recommendation system for repetitively purchasing items in E-commerce," 2014 Int. Conf. Big Data Smart Comput. BIGCOMP 2014, pp. 96–98, 2014.
- [16] Y. Fan, Y. Shen, and J. Mai, "Study of the model of e-commerce personalized recommendation system based on data mining," *Proc. Int. Symp. Electron. Commer. Secur. ISECS 2008*, pp. 647–651, 2008.
- [17] M. I. Martín-Vicente, A. Gil-Solla, M. Ramos-Cabrer, Y. Blanco-Fernández, and M. López-Nores, "Avoiding fake neighborhoods in e-commerce collaborative recommender systems: A semantic approach," *SMAF '09 - 4th Int. Work. Semant. Media Adapt. Pers.*, pp. 9–14, 2009.
- [18] Y. Luo, "An analysis of web mining-based recommender systems for E-commerce," *Proc. 2012 Int. Conf. Comput. Appl. Syst. Model. ICCASM 2012*, pp. 0167–0170, 2012.
- [19] F. Huseynov, S. Y. Huseynov, and S. Özkan, "The influence of knowledge-based e-commerce product recommender agents on online consumer decision-making," *Inf. Dev.*, vol. 32, no. 1, pp. 81–90, 2016.
- [20] P. Gatchalee and Z. Li, "Ontology Development for SMEs E-commerce website based on Content Analysis and its Recommendation System," pp. 7–12, 2013.
- [21] C. Chen, D. Wang, and Y. Ding, "User actions and timestamp based personalized recommendation for E-commerce system," *Proc. - 2016 16th IEEE Int. Conf. Comput. Inf. Technol. CIT 2016, 2016 6th Int. Symp. Cloud Serv. Comput. IEEE SC2 2016 2016 Int. Symp. Secur. Priv. Soc. Netwo*, pp. 183–189, 2017.
- [22] J. F. Miao, "Design and implementation of personalization recommendation system in mobile e-commerce," *Adv. Mater. Res.*, vol. 989–994, pp. 4506–4509, 2014.
- [23] M. Aprilianti, R. Mahendra, and I. Budi, "Implementation of weighted parallel hybrid recommender systems for e-commerce in Indonesia," *2016 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2016*, pp. 321–326, 2017.
- [24] F. Dong, J. Luo, X. Zhu, Y. Wang, and J. Shen, "A personalized hybrid recommendation system oriented to E-commerce mass data in the cloud," *Proc. - 2013 IEEE Int. Conf. Syst. Man, Cybern. SMC 2013*, pp. 1020–1025, 2013.
- [25] A. Gustafsson, F. Ekdahl, and B. Bergman, "Conjoint analysis: A useful tool in the design process," *Total Qual. Manag.*, vol. 10, no. 3, pp. 327–343, 1999.
- [26] J. Zhang and P. Pu, "Survey of Solving Multi-Attribute Decision Problems," *EPFL Tech. Rep. No IC/2004/54*, no. July 2010, pp. 1–14, 2004.