Using factor analysis and regression analysis to develop mathematic modeling to support online application implementation in transportation business

W Sardjono¹, E Selviyanti², W G Perdana³

¹Information Systems Management Department, Bina Nusantara University, Jakarta, Indonesia
²Health Department, PoliteknikNegeriJember, Jember, Jawa Timur, Indonesia
³Post Graduate Program of Environmental Science, School of Environmental Science

Email: wahyu.s@binus.ac.id, ernaselviyanti@polije.ac.id, widhilaga.gia@ui.ac.id

Abstract. Factor analysis is used to reduce data or summarize, from the old variable which is widely changed to a few new variables called factors, and still contains most of the information contained in the original variable. While regression analysis studies the relationship between one or more variables / independent variables (X) with one dependent variable (Y). The purpose of this paper is to describe how the factor analysis method can be used to provide solutions to operational problems such as problems with public transportation services with online applications, for this reason the research was conducted in identifying what factors influence the quality of online applications, and henceforth can be used in building mathematical models that can explain practical applications and predictions in the future. This research is explained in a quantitative descriptive form using exploratory factor analysis and regression analysis methods to process primary data and build models. The conclusion of this study is that there are 22 indicators and three factors that determine the quality of the application. The three factors produced are system and information factors, user impact factors, and promotion factors. These three factors can henceforth become a model in determining customer satisfaction as an end user of public transportation with an online application.

1. Introduction

About 5 (five) years ago, a startup company presented online applications for public transportation and as an intermediary that connects public transport drivers with customers who want to use transportation facilities with online applications [1], and only a few months later after the online application was launched, the online application was found to have been downloaded by more of 100,000 downloaders. In line with the development of mobile phones in almost all major cities in Indonesia, the development of online applications is also increasingly widespread development [2]. So that demands for services to providers and managers of online application services are also increasing, especially the quality of online applications in serving their customers. [3] on the quality of the website of the Republic of Indonesia Education and Training Center proving that the availability of online applications for quality public transportation can increase the satisfaction of users of these transportation facilities [4]. This research has resulted in several new factors, which can subsequently be used to build models that can explain practically the quality of online applications from the point of view of the customer as a user of public transportation using an online application[5]. The results of this study are expected to help public transportation companies with online applications in measuring product quality and its effect on customer satisfaction.

2. Methodology

The study was conducted by collecting data from already existent information, through secondary sources and from original sources, for example primary data [6]. The secondary data was collected from already published articles, reports, books, book sections, as well as from websites [7]. This was
done through a literature review on the concepts of knowledge, knowledge management, and policy process and online application adaptation [8]. This method of data collection was chosen for the reason that such a study needed a theoretical framework in place before conducting any form of original research [9]. A literature review was needed for the comparison of frameworks undertaken to show that a commercial concept like online application has social benefits and to determine if there were any gaps in the adaptation online application platform framework [10]. The primary data was collected through communication, which took place via a self-administered questionnaire. The questionnaire was chosen to cover the large group of participants in a short span of time and avoid any data [11]. This method was chosen over the interview method to give the participants more freedom to respond. The questions were close-ended chosen due to the non-personal nature of the questionnaire [12].

Quality dimensions of usability:
1. Easy to operate ($X_{11}$).
2. User interaction with the application, fast and easy ($X_{12}$).
3. Easy to read navigation ($X_{13}$).
4. Promotion will be made of the application intensively ($X_{14}$).
5. Attractive application display ($X_{15}$).
6. Appropriate design or layout ($X_{16}$).
7. Display applications according to e-commerce application types ($X_{17}$).
8. Creates a positive experience ($X_{18}$).

The dimension of information quality:
1. Accurate information ($X_{21}$).
2. Reliable information ($X_{22}$).
3. Information in a timely manner ($X_{23}$).
4. The application provides relevant information ($X_{24}$).
5. Information that can be understood ($X_{25}$).
6. Information according to user requirements ($X_{26}$).
7. Information in the right format ($X_{27}$).

Quality dimensions of interaction and service:
1. Reputation ($X_{31}$).
2. Security system ($X_{32}$).
3. Personal information ($X_{33}$).
4. Good feeling of personalization ($X_{34}$).
5. A good community atmosphere ($X_{35}$).
6. Easy to communicate online ($X_{36}$).
7. Promised services ($X_{37}$).

The fixed influence model was selected using all the subset regression analyzes in three, four, and five experimental factors for each of the two types of proppants. The five-factor model supported is the most useful model for both types of proppants from a predictive point of view [12]. The five-factor prediction model is also compared to the supported fracture conductivity model and the publication of different reports published in the literature [13]. End-user expectations of an application for the availability of an application that is also an additional independent variable in this study so that the total indicator of the independent variable becomes 25 indicators, namely:
1. Fast application response time ($X_{41}$).
2. Applications can be accessed at any time ($X_{42}$).
3. Applications are easy to find in the application store ($X_{43}$).

In this research, the type of research used is explanatory research or explanatory level research. Based on the type of explanatory research or explanatory level research, the type of research is associative research. Associative research is research that aims to determine the relationship between two or more variables. Population is a generalization area consisting of objects / subjects that have certain qualities and characteristics determined by researchers to be studied and then drawn conclusions.

1. Using a mobile device before.
2. Never accessed the public transportation with online application before.
3. Never used public transportation with online application services before.

3. Result
From the respondent's answer, regardless of gender differences, it appears that the respondent's work. The results of data processing using factor analysis. Furthermore, in table 2, the results of the regression analysis will be shown at a significance level of 0.05.

| Item | Starting Conditions | Factor analysis result |
|------|---------------------|------------------------|
| Number of Factors | 4 | 3 |
| Number of Indicators | Factor number -1 = 8 | Factor number -1 = 14 |
| | Factor number -2 = 7 | Factor number -2 = 8 |
| | Factor number -3 = 7 | Factor number -3 = 1 |
| | Factor number -4 = 3 | Total Indicator = 23 |

| Variable | Coefisien | t count | t tabel | Remark |
|----------|-----------|---------|---------|--------|
| Constant | 5.40      |         |         |        |
| $X_1$    | 0.45      | 3.90    | 1.66    | Significant |
| $X_2$    | -0.43     | -0.28   | 1.66    | Un-Significant |
| $X_3$    | 0.01      | 0.05    | 1.66    | Un-Significant |
| F count  | = 5.10    |         |         |        |
| F table  | = 2.68 < F table< 2.76 |         |         |        |

4. Discussion
Data processing using the factor analysis method in this study obtained a number of indicators that cluster in 3 (three) new independent variables that describe the conditions of using online applications to support operational facilities for public transportation. Three new variables and the results of grouping the indicators are:

The first variable ($X_1$) consists and namely:
1. Simplicity
2. User orientation, fast and easy.
3. Easy to read navigation.
4. Build a positive experience.
5. Accurate information.
6. Reliable information.
7. Information in a timely manner.
8. Relevant information.
9. Clear information.
10. User requirement achievement.
11. Simple information format.
12. Reputation.
13. Services promised.
14. Fast response time.

The new factor of information system is the first factor that represents a number of indicators which are grouped above.

The second variable (X₂) consists and namely:

1. Attractive application display
2. Appropriate layout
3. Display applications according to e-commerce application types.
4. System security.
5. Personal information.
6. Good personalization feeling.
7. Good community atmosphere.
8. Easy to online communication

The new factor of customer un-convenience is the second factor that represents a number of indicators which are grouped above.

The third variable (X₃) consists of one indicator, namely the promotion of the existence of an intensive application.

The new factor of promotion is the third factor that represents a number of indicators which are grouped above.

Based on the results of data processing and then the results of the regression analysis process, can be formulated into the predictive mathematical model obtained, namely:

\[ Y = 5.40 + 0.45 X_1 - 0.43 X_2 + 0.01 X_3 \]

\( Y \) = The satisfaction of mobile application end users
\( X_1 \) = System information (between -3.24 to 1.70)
\( X_2 \) = Customer un-convenience (between -2.33 to 2.07)
\( X_3 \) = Promotion (between -2.86 to 1.72)

The results of the calculation of the regression analysis in this study also proved to be significant and not biased after passing all the assumption tests. Model simulation results are presented in table 3.
### Table 3. Result of Model simulation

| Condition     | C | X₁ | X₂ | X₃ | Y   |
|---------------|---|----|----|----|-----|
| Current       | 5.40 | 0  | 0  | 0  | 8.28 |
| Expected      | 5.40 | 1.70 | -2.33 | 1.72 | 9.68 |
| Un-Expected   | 5.40 | -3.24 | 2.07 | -2.86 | 6.02 |

Based on the results of calculations in the method of regression analysis and with factor score data which is the output of the data processing process with factor analysis that has been done before, obtained the average level of customer satisfaction today is 8.396 on a scale between 1 and 10. By paying more attention to information systems factors, effects on users, and promotions, companies can increase the level of customer satisfaction up to 9.68. Conversely, the level of customer satisfaction can drop to 6.02 if the company does not pay attention to these three factors.

5. Conclusion

From the results of the discussion related to the use of factor analysis and regression analysis to determine mathematical models in supporting the implementation of online applications, in this example Go-Jek is taken as a business actor in the field of transportation, it can be concluded:

1. Factors that can influence people's decisions to use public transportation with online applications are information system factors (X₁), customer un-convenience factors (X₂), and promotion factors (X₃).

2. Mathematic Equation models finding:

\[ Y = 5.40 + 0.45X_1 - 0.43X_2 + 0.01X_3 \]

With the minimum and maximum values:

-3.24 \( \leq \) \( X_1 \) \( \leq \) 1.70
-2.33 \( \leq \) \( X_2 \) \( \leq \) 2.07
-2.86 \( \leq \) \( X_3 \) \( \leq \) 1.72

3. Research Limitations to accompany the respondents, this can cause inaccuracies in filling out the questionnaire, such as inaccurate answers, respondents filled the questionnaire less seriously and did not focus on the statements submitted, as well as differences in understanding of statements between researchers and respondents.

4. Managerial Implications

1. Functionally, existing online applications need to be developed, both in terms of the format and substance contained in them, so that they become easier and more attractive to users.

2. Motorcycle taxi drivers who are partners of the company need to be increased so that customers do not need a long time to get GO-JEK services. Nevertheless, the process of receiving, guiding and supervising motorbike riders must still be considered.

Acknowledgments

Thank you to Bina Nusantara University for giving us the opportunity and support to conduct research.

References

[1] CoppolaP and Silvestri F 2018 Estimating and visualizing perceived accessibility to transportation and urban facilities Transportation Research Procedia 31 136

[2] Alexey L, Alexander M and Maksim S 2018 Special generators in tasks of transportation demand assessment. Transportation Research procedia 36 434
[3] Alexey L 2017 Application of geoinformation technologies for the transportation demand estimation Transportation Research Procedia 20 406

[4] Aaron C, Alireza A, Hanjin H and Stuart SC 2018 Building information modeling (BIM) for transportation infrastructure – Literature review, applications, challenges, and recommendations Automation in construction, 94 257

[5] Yanli W, Sabyasachee M, Xin Y, Linbo L and Bing W 2017 The Application of integrated multimodal metropolitan transportation model in urban redevelopment for developing countries Transportation Research Procedia 25 2990

[6] Andreas K, Ronald M and Alexander S 2016 The effects of collecting and connecting activities on knowledge creation in organizations The Journal of Strategic Information Systems 25 243

[7] Medyawati H and Mabruri A 2012 Website quality: Case study on local government bank and state own bank in bekasi city Procedia - Social and Behavioral Sciences, 65 1086

[8] T. Vivekanandan T and Narayanan S J 2019 A Hybrid Risk Assessment Model for Cardiovascular Disease Using Cox Regression Analysis and a 2-means clustering algorithm Computers in Biology and Medicine 113

[9] Bilous I, Deshko V and Sukhodub I 2018 Parametric analysis of external and internal factors influence on building energy performance using non-linear multivariate regression models Journal of Building Engineering 20 327

[10] H. Kashani Rad H, Shariatmadar H and Ghalehnovi M 2020 Simplification through regression analysis on the dynamic response of plates with arbitrary boundary conditions excited by moving inertia load Applied Mathematical Modelling 79 594

[11] Peigen Yu P, Low, M Y and Zhou W 2018 Design of experiments and regression modelling in food flavour and sensory analysis: A review Trends in Food Science & Technology 71 202

[12] Gkioulekas I and Papageorgiou L G 2018 Piecewise Regression through the Akaike Information Criterion using Mathematical Programming IFAC-Papers On Line 51 730

[13] SAI-Salem S M, Al-NasserA and Al-Dhafeeri A T 2018 Multi-variable regression analysis for the solid waste generation in the State of Kuwait Process Safety and Environmental Protection 119 172