yönetim ve ekonomi araştırmaları dergisi
journal of management and economics research
Cilt/Volume: 17 Sayı/Issue: 4 Aralık/December 2019 ss./pp. 269-280
İ. Topal Doi: http://dx.doi.org/10.11611/yead.542249

ESTIMATION OF ONLINE PURCHASING INTENTION USING DECISION TREE

Dr. İbrahim TOPAL *

ABSTRACT

It is very difficult to estimate consumer behavior due to different variables. There are also differences between the online consumer and the traditional ones. While there are studies for the prediction of purchasing behavior of online consumers, there is need for further studies with larger data including different features. Large data is difficult to obtain due to restrictions on private information and causes the analysis systems run for a long time. So, in this study, it is aimed to create a meaningful rule by estimating the purchasing behavior of online consumers with fewer data. After selecting the Fisher Score feature in a current and open database, training and test data were determined with K fold and a rule was created with Decision Tree. As a result, it can be suggested that it is possible to determine the purchasing behavior of online consumers with high accuracy by using a single feature.

Key Words: *Online Purchasing Intention, Artificial Intelligence, Fisher Feature Score Selection, Decision Tree, K-Fold.*

JEL Codes: *M30.*

KARAR AĞACI KULLANARAK ÇEVRİMİÇİ SATIN ALMA NİYETİNİN TAHMİNİ

ÖZET

Tüketici davranışlarını tahmin etmek çok fazla değişkene bağlı olması nedeniyle oldukça zordur. Aynı zamanda çevrimiçi tüketiciyle geleneksel tüketici arasında farklılıklar bulunmaktadır. Online tüketicilerin satın alma davranışını tahmine yönelik bir süredir çalışmalar olmakla birlikte çok sayıda özelliğe sahip büyük verilere ihtiyaç duyulmaktadır. Büyük verilerin, kişisel bilgilere yönelik kısıtlamalar nedeniyle elde edilmesi zor olmakta ve analiz sistemlerini uzun süre çalışmasına sebep olmaktadır. Bu bağlamda, çalışmada online tüketicilerin satın alma davranışını daha az veriyle tahmin ederek anlamlı bir kural oluşturmak amaçlanmıştır. Güncel ve açık bir veri tabanında Fisher skor özellik seçme yapıldıktan sonra K fold ile eğitim ve test verileri belirlenerek karar ağacı ile kural oluşturulmuştur. Sonuç olarak tek bir özellik kullanılarak çevrimiçi tüketicinin satın alma davranışının yüksek doğruluk oranıyla tespitinin mümkün olduğu görülmüştür.

* National Defense University, Yalova, Turkey, e-mail: itopal@msu.edu.tr

Makale Geçmişi/Article History

Başvuru Tarihi / Date of Application	: 21 Mart / March 2019
Düzeltme Tarihi / Revision Date	: 12 Kasım / November 2019
Kabul Tarihi / Acceptance Date	: 30 Aralık / December 2019

Araştırma Makalesi/Research Article

Anahtar Kelimeler: *Çevrimiçi Satın Alma Niyeti, Yapay Zeka, Fisher Skor Özellik Seçme, Karar Ağacı, K-Katlamalı.*

JEL Kodları: *M30.*

1. INTRODUCTION

With the widespread access of the Internet, it is possible to reach information and other people quickly and easily and this brought various opportunities together. After 1990's, online shopping has increased and allowed to the creation of an almost perfect market where products can be compared globally (Kuttner, 1998).Consumers not only have access to technical or experience-based information about the products or services they need, but they can also purchase these products and services. Today, many products from travelling to grocery shopping can be purchased online. Online shopping is rapidly increasing and market volume is expanding. Global retail sales on the Internet was reported to have risen from $ 1.3 trillion in 2014 and by 2019 it's expected that the sales will reach to $ 3.5 trillion and by 2021 to $ 4.9 billion. Accurate estimation of consumers' purchase behavior in the growing market will be able to offer various opportunities to businesses.

When consumers need a product or a service, they collect information, evaluate alternatives, decide and purchase. Today's consumers can use the internet at every stage of this process (Kotler & Lee, 2007). For businesses that want to sell online, selling products and services globally is an advantage while there is a disadvantage of more competition. This makes it difficult to predict the purchasing behavior of consumers who may influence the short or long term plans of businesses in various fields.

Sales estimation can affect businesses from raw material supply to advertising costs and payments. This estimation can be made based on experience or analysis in the market. Thanks to technological developments devices that emerged as a result of artificial intelligence have become one of the analysis methods. From the 1980s to the present, businesses have been using artificial intelligence in fields such as finance, advertising, and sales estimation(Wong, Bodnovich, & Selvi, 1996). Estimation of purchasing behavior, which has become an issue that can affect the success of businesses, in general, can be realized with large data and artificial intelligence methods.

Estimates can be made to benefit businesses with a large number of data on consumer transactions in internet shopping. However, one of the most important problems here is certain information is considered private and cannot be used (Park & Huh, 2019), the second is evaluating a large number of data takes time. Simplifying rules to make instant predictions in online purchases or forwarding instant promotions to the right person will reduce costs. In this context, decision trees are one of the methods that can be used. Decision trees, one of the multistage decision making, have been widely used in several disiplines because of easy use, free ambiguity, and robust (Song & Lu, 2015). By using decision trees classification, it is possible to derive rules which are feasible and are for purchasing online shopping sites.

In this study, it is aimed to estimate successful consumer purchasing behavior in internet shopping by using fewer features. An update and active database, Fisher feature selection, k-Fold cross-validation, and decision trees were used for analysis.

2. LITERATURE REVIEW

In terms of business managers, knowing how much the product will be transformed into money is effective in many areas from the current account balance to the balance of payments. Sales estimates that support decision makers are important for in achieving information about the goals of the business. One of the analysis methods to be used in estimating the purchase is artificial intelligence. The studies of Dutta (1994) and Nam(1995) are among the first examples in which the orders are estimated by neural networks. In the following period, many studies have been carried out to estimate consumers' purchasing behavior.

Vellido (2000), in order to estimate consumers' online purchasing behavior, mentioned that consumers' web shopping was collected in four factors. These are product perception, shopping experience, customer service, and consumer risk. 44 items were selected from the open publicly database for the use in analysis. And performance was tested with two different methods as logistic discriminant and neural network. As a result of the study, it was found out that the neural networks give better results (Vellido, Lisboa, & Meehan, 2015).

Another effect of online purchase estimating is on pricing. Pricing is one of the four main elements of marketing, known as 4P (product, promotion, price, and place). Pricing directly affects sales, profit rate, and competition. The ideal price can be found out by estimating the instant purchase for online consumers. In the study of Gupta (2014), visits from different data sources, visitor characteristics, purchasing history, web data, and content perception were used. Data mining, big data technologies, and logistic regression were used in the study. As a result of the study, it is seen that thanks to purchase estimation more accurate pricing could be made (Gupta & Pathak, 2014). The purchasing estimation can be useful in other areas besides the pricing. However, being aware of the fact that online consumers are different from traditional consumers will contribute to the development of an estimation method.

The estimation of the online consumer's purchasing behavior and the procedures performed accordingly are different from the traditional consumers. This is a good example of electronic word of mouth (e-wom) communication, defined as the communication between consumers on the internet without commercial concern about products or services. E-wom, which is an important component of online shopping, differs from face to face communication in various ways in transmitting the information of the product. The transfer of experiences on the internet is different from the communication in daily life in terms of speed, number of people reached, and persistent knowledge (Hoffman & Novak, 1996). At the same time, consumers rely heavily on comments on the internet (Nielsen, 2013). In this context, it is possible to estimate the purchasing behavior based on the comments of the consumers. In Qiu, Lin and Li (2015)'s study, it was successful model in estimating the purchase of the COREL (Customer Purchase Prediction Model) model, which was generated by consumers' evaluations (number of comments, average score, product-shelf-date, the most recent review) and price, brand information. In the analysis of this model, Bayesian discrete choice model was used (Qiu, Lin, & Li, 2015). The variety of models in the estimation of consumer purchasing results from the interest of different units in the enterprise. One of these units is customer relationship management (CRM)

CRM has an important role in understanding the behavior of consumers and maintaining a long and healthy relationship. For CRM to be successful, there is a need for serious data bases and data mining for consumer behavior. The data mining models used by CRM are association, classification, clustering, forecasting, regressions, sequencing discovery, and visualization. Gordini (2014) is a detailed data mining on CRM. In the study, 42 types of data were used and decision tree, one of the machine learning methods, was estimated. As a result of the study, the customers were divided into eight different categories according to the probability of purchasing between 1% and 92% (Gordini, Sanpaolo, & Veglio, 2015).

Another study, which made a considerable prediction, was conducted by Sakar et al. (2018). At the same time, the intention of online consumers to purchase is predicted by data mining and artificial intelligence methods. 18 properties and 12,330 session information of the consumers were analyzed with C4.5, random forest, support vector machine and multilayer perception classifier methods and successfully estimated 87.24%(Sakar, Polat, Katircioglu, & Kastro, 2018).

Data mining is the analysis and exploration of large databases to find very important information that can help decision makers. In a study on car purchasing Wah, Ismail, and Fog (2011) 1935 cases using data mining. On decision treees, CART algorithm, logistic regression (LR), and artificial neural networks were used and proved 89% accurate. The highest sensivity value was achieved in logistic regression.

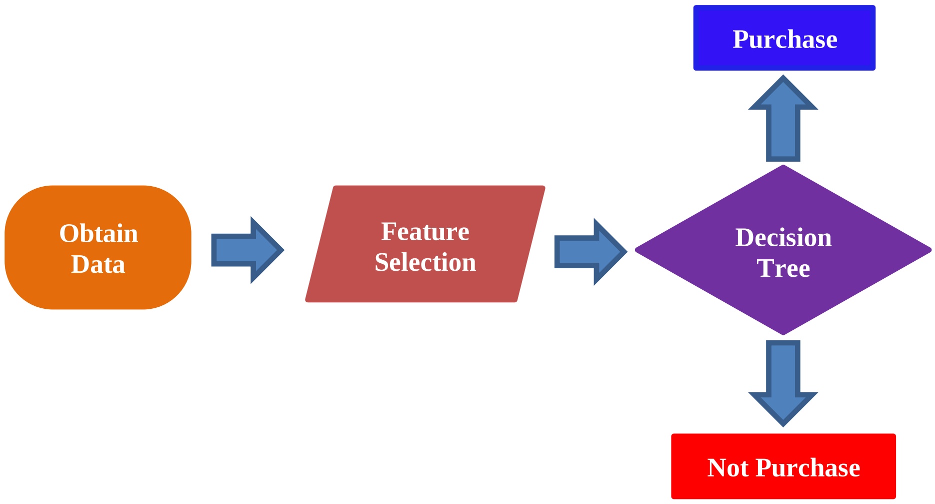
In a study by Fokin and Hagrot (2016), the aim was to predict user behavior with Bestbuy data made available to public for conducting data mining research. In the study, decision tree ID3 algorithm was used. The product title and description was shown to be two important aspects of product innovation and the study issued a tree diagram applicable to all products.

It is seen that sales estimations, which are frequently needed by different units of enterprises, are estimated by various methods but they need big data as a common feature. Due to the fact that customers can leave the page in seconds, successful estimations with less data make it easier to obtain data and shorten process of artificial intelligence transaction.

3. METHODOLOGY:

The process flow diagram of the operations is shown in Figure 1. According to the diagram, firstly the data were taken for use in artificial intelligence algorithm. The second step was the Fisher feature selection process. The study is expected to contribute to the development of sales with rules that can be used by businesses. Therefore finally, in the literature, the decision trees classification method which is simple and frequently used and strong in rulemaking is used. The rule that people will purchase / not purchase with the decision tree algorithm that is run with K Fold cross-validation has been revealed. In the study Matlab R2018a software was used.

Figure 1. The Flow Diagram for Predicting Online Purchase



3.1. Data

In the study, data which has been used, belong to the research on online purchasing behavior (Sakar et al., 2018) which was, on the open database website UCI Machine Learning Repository. These data pertaining to a commercial web page was derived from Google Analytics. Data covering an annual period and the impact of special campaigns or days are avoided. The purchase intention is kept in the database with binary classification, either true or false. In the database consisting of 12,330 sessions, there is no shopping in 10,422 and 1,908 in shopping. In the database consisting of 12330 sessions, revenue value indicates shopping transaction is negative in 10422 records and positive in 1908 records. The purchase/non - purchase rate is approximately 85% - 15%. It is observed that the purchasing intention is imbalanced. It is common for online consumers to not make a purchase of any product during most visits to their shopping sites. The data and explanations are given in Table 1.

Table 1. Features and Descriptions

No:	Feature Name	Feature Description
1.	Administrative	Pages' number visited on account management
2.	Administrative Duration	The total time spent on account management pages (in seconds)
3.	Informational	Pages' number visited on shopping informational pages (communication and address)
4.	Informational duration	The total time spent on informational pages (in seconds)

5.	Product related	Pages' number visited on product related
6.	Product related duration	The total time spent on product related pages (in seconds)
7.	Bounce rate	The pages' mean bounce rate value
8.	Exit rate	The pages' mean exit rate value
9.	Page value	Mean page value
10.	Special day	Closeness of visiting time to a special day
11.	Operating Systems	Operating System
12.	Browser	Browser type
13.	Region	Geographic region
14.	Traffic Type	Visitor's reach type to Web site (e.g., banner, direct)
15.	Visitor Type	New Visitor, Returning Visitor, Other
16.	Weekend	Weekend or not
17.	Month	Month value
18.	Revenue	The visit has been finalized with a transaction or not

Reference : Sakar, C. O., Polat, S. O., Katircioglu, M., & Kastro, Y. (2018). Real-time prediction of online shoppers ' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. *Neural Computing and Applications, 0*. https://doi.org/10.1007/s00521-018-3523-0

3.2. Fisher Score Feature Selection:

In artificial intelligence algorithms, a large number of properties have a negative effect on the process of obtaining and processing data. Therefore, a pre-treatment is performed to reduce the number of features and to facilitate the work of the classifier. These processes, which enable the extraction of non-classification properties, are called feature selection methods.

Generally, feature selection methods are examined in three categories: filter-based, wrapper-based and embedded methods (Guyon & Elisseeff, 2003). Filter-based methods are sort of properties of the learning algorithm from the highest to the lowest as pre-process steps. Wrapper-based methods work with the properties of the learning algorithm to be used. Embedded methods combine feature selection and learning algorithms (Gu, Li, & Han, 2012).

Filtration methods are an important methods in determining the most suitable sub-spaces of the class, regardless of the classifiers based on statistical methods(Huerta, Duval, & Hao, 2010). One of the filtering methods is the fisher score. This method calculates a relationship score using the mean and standard deviation values of feature for each class. Fisher Score calculation formula is shown in equation (1) (Bolón-Canedo, Sánchez-Marono, Alonso-Betanzos, Benitez, & Herrera, 2014; Budak, 2018).

$$F(X_i) = \frac{|\mu_i^+ - \mu_i^-|}{\sigma_i^+ - \sigma_i^-} \qquad \textbf{Eq.(1)}$$

In Equation 1., It is indicated that + and – is different classes for a two-class problem, μ_{i+} and μ_{i-} is average of the classes, σ_{i+} and σ_{i-} is standard deviation of the classes. With this method, the feature selection process is made by selecting the desired numbers of features starting from the top order after the properties are sorted from big to small according to the calculated scores.

3.3. Decision Tree:

A decision tree is a algorithm that makes systematic analysis to extract valuable rules and relationships from a data set containing a large number of records and is often used in classification or prediction (Lee, Lee, & Park, 2007). Decision tree, which is one of the most commonly used machine learning methods, is used for generating understable rules, creating rules without requiring too much action, using with continuous and categorical variables, and easy indication.

There are two different methods in decision trees: classification and regression (CRT). In the classification method, the decision tree looks like a top-down tree. Each node extending up to the leaves in the tree is formed by decisions resulting from processing of the data. In the CRT method, the tree model can be grouped according to variables. Variables with the best distinction are included in the nodes. Nodes from the root to the leaves are repeated with the same or different variables. In order to minimize group variances, a rule based decision tree is created. If the class label of the data is categorical, the classification tree is used, if it is a continuous numerical value, CRT is used(Ucar & Topal, 2018).

In this study, "Revenue", class label, is categorical. Therefore, the classification tree method was used. Besides with, k-fold cross-validation method was used because of imbalanced data.

3.4. K Fold Cross Validation:

K Fold cross validation method is one of the holdout methods. The data set is divided into subsets as much as the given number of k and the process is repeated as k. In each repetition, k subset is divided into test and k-1 training and the hypothesis is tested. The default accuracy value is found in each cross-validation layer by dividing the data in random form. At the end of the process, the default accuracy is found by taking the average of all the k trials (Kohavi, 1995) . In this study, because of the low bias and variance, cross-validation 10 fold cross-validation was used(Han, Kamber, & Pei, 2014)..

4. FINDINGS:

With the Fisher Features Selection Algorithm, the properties that have low classification power have been subtracted and the accuracy value is calculated. The process was continued as long as there was no major negative change in accuracy. As a result, it is seen that 16 characteristics can be deduced and the same performance can be predicted in a single feature. It is seen that the most powerful feature in the classification is “Page Value".

Google Analytics describes Page Value as follows; “the average value for a page that a user visited before landing on the goal page or completing an Ecommerce transaction (or both)”(“How Page Value is calculated," 2019). In this study, Page Value is between 0 and 361.76. Since Page Value has a different value of 2707, a general rule cannot be adopted at the end of the decision tree.

A new procedure for Page Value has been developed to make a meaningful rule. The average of non-zero values is 26.59, which is about 25. Therefore, new values are given by using 25 and multiples. In the light of the new values, Page Value 0-25 was assigned as 1, 26-50 as 2, 51-75 as 3 … and 351-375 as15. Fisher score feature selection and decision tree operations were repeated with new data. The feature selection process is identical to the previous one, and the performance values are acceptable because they are above 80%. The highest accuracy value is reached only if the Page Value feature is processed. The results are given in Table 2.

Table 2. Fisher Feature Selection Results

Fisher P.N.	Features	Accuracy-1	Accuracy -2*
1.	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	0,8676	0,8223
2.	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17	0,8649	0,8246
3.	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 16, 17	0,8663	0,8194
4.	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 16, 17	0,8679	0,8221
5.	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 16, 17	0,8626	0,8238
6.	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 16	0,8662	0,8185
7.	1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 16	0,8659	0,8247
8.	1, 2, 3, 5, 6, 7, 8, 9, 11, 16	0,8632	0,8235
9.	1, 3, 5, 6, 7, 8, 9, 11, 16	0,8691	0,8206
10.	1, 3, 5, 6, 7, 8, 9, 16	0,8649	0,8224
11.	1, 3, 5, 6, 7, 8, 9	0,871	0,8254
12.	3, 5, 6, 7, 8, 9	0,8611	0,8146
13.	3, 5, 7, 8, 9	0,8611	0,818
14.	3, 5, 7, 9	0,8655	0,8233
15.	3, 7, 9	0,8672	0,836
16.	7, 9	0,8667	0,8477
17.	9	0,873	0,8839

Fisher P.N.= Fisher Score Feature Selection Process Number
*Accuracy-2= Results after "Page Value" is grouped as 25 and multiples

The rule-based decision tree image is given in Figure 3. Accordingly, the tree controls the value of "Page Value" from the data set. If the value of Page Value is equal to or greater than 1.5, it is possible to make a purchase. If the value is less than 1.5, it is estimated that the purchase will not take place. The code of the decision tree is shown below.

The results after the feature selection are shown in Table 3. Accuracy value is 88,39 true positive 74,71% and true negative value 89,95% is good fit for social sciences. The AUC and Kappa values are mean fit and the F-Score value is acceptable.

Table 3. Classification Results

Accuracy	AUC	Kappa	F Score	True Positive Rate	True Negative Rate
88,39	0,67	0,44	0,54	74,71	89,55

Decision Tree Code for Purchasing Intention

True: Class 1, False: Class 2

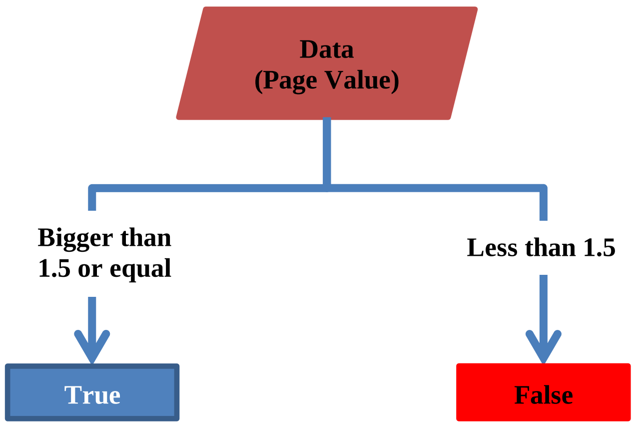
Decision tree for classification

1 if x1<1.5 then node 2 elseif x1>=1.5 then node 3 else 2

2 class = 2

3 class = 1

Figure 2. Decision Tree for Purchasing Intention-2



5. CONCLUSION:

Sales figures are one of the most important data which can affect many units such as raw material procurement or new product development. In the estimation of the purchase behavior, it is possible to examine today's consumers in two types, online and traditional. Although online purchasing increase rapidly, problems appear due to reasons such as the protection laws of personal information in front of owning and using this information. Therefore, it is necessary to reach the most successful result with the least data in the estimation of the purchase.

In this study, purchase behavior has been estimated by using the open database. With the Fisher feature selection method, it is seen that the decision tree can be extracted with a single feature by maintaining the accuracy value at acceptable levels. This feature is related to the page value and has an accuracy value of over 80%.

It is observed that there are studies, with %60 accuracy level in the business management literature (Atsalakis, Atsalaki, & Zopounidis, 2018; King, Abrahams, & Ragsdale, 2014). However, the acceptable accuracy level is 80% in studies affecting human health, such as medicine (Reha Alpar, 2016). Although this study is about business management, it has 88% accuracy rate. In addition to this, Page Value has been revalidated according to a procedure in order to infer a meaningful rule from the study.

The result of the analysis suggests that, if the Page Value is 1.5 and above, the transaction for the purchase can be started. It is possible for this business to offer instant opportunities and encourage more shopping for customers, whose Page Value is $37.5 or more.

Page Value is the average of the target page and the page before it reaches. If the previous page is not an e-commerce page, the value of the page is zero. In other words, if the sales page is accessed from a valuable e-commerce page, the purchase is affected. In this context, it can be concluded that the companies that want to sell on the internet would benefit from advertising on other sales pages. For example, shoe or bag ads on a page selling clothes; swimsuit ads on a page selling travel tour; furniture ads on a page selling TV can contribute to their sales.

REFERENCES

Atsalakis, G. S., Atsalaki, I. G., & Zopounidis, C. (2018). Forecasting the success of a new tourism service by a neuro-fuzzy technique. *European Journal of Operational Research*, *268*(2), 716–727. https://doi.org/10.1016/J.EJOR.2018.01.044

Bolón-Canedo, V., Sánchez-Marono, N., Alonso-Betanzos, A., Benitez, J. M., & Herrera, F. (2014). A review of microarray datasets and applied feature selection methods. *Information Sciences*, *282*, 111–135.

Budak, H. (2018). Özellik Seçim Yöntemleri ve Yeni Bir Yaklaşım. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, *22*(Özel Sayı), 21–31. https://doi.org/10.19113/sdufbed.01653

Dutta, S., Shekhar, S., & Wong, W. Y. (1994). Decision support in non-conservative domains: Generalization with neural networks. *Decision Support Systems*, *11*(5), 527–544. https://doi.org/10.1016/0167-9236(94)90023-X

Gordini, N., Sanpaolo, I., & Veglio, V. (2015). *Customer relationship management and data mining : A classification decision tree to predict customer purchasing behavior in global market*. https://doi.org/10.4018/978-1-4666-4450-2.ch001

Gu, Q., Li, Z., & Han, J. (2012). Generalized fisher score for feature selection. *ArXiv Preprint ArXiv:1202.3725*.

Gupta, R., & Pathak, C. (2014). A Machine Learning Framework for Predicting Purchase by online customers based on Dynamic Pricing. *Procedia - Procedia Computer Science*, *36*, 599–605. https://doi.org/10.1016/j.procs.2014.09.060

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3*(Mar), 1157–1182.

Han, J., Kamber, M., & Pei, J. (2014). *Data mining: Data mining concepts and techniques*. *Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013*. https://doi.org/10.1109/ICMIRA.2013.45

Hoffman, D. L., & Novak, T. P. (1996). Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations. *Journal of Marketing*, *60*(3), 50. https://doi.org/10.2307/1251841

How Page Value is calculated. (2019). Retrieved March 12, 2019, from https://support.google.com/analytics/answer/2695658?hl=en

Huerta, E. B., Duval, B., & Hao, J.-K. (2010). A hybrid LDA and genetic algorithm for gene selection and classification of microarray data. *Neurocomputing*, *73*(13–15), 2375–2383.

King, M. A., Abrahams, A. S., & Ragsdale, C. T. (2014). Ensemble methods for advanced skier days prediction. *Expert Systems with Applications*, *41*(4 PART 1), 1176–1188. https://doi.org/10.1016/j.eswa.2013.08.002

Kohavi, R. (1995). The Power of Decision Tables. *ECML*.

Kotler, P., & Lee, N. (2007). *Kamu Sektöründe Pazarlama*. İstanbul: Mediacat.

Kuttner, R. (1998). The net: a market too perfect for profits. *Business Week*, *3577*(1).

Lee, S., Lee, S., & Park, Y. (2007). A prediction model for success of services in e-commerce using decision tree : E-customer ' s attitude towards online service, *33*, 572–581. https://doi.org/10.1016/j.eswa.2006.06.005

Nam, K., & Schaefer, T. (1995). Forecasting international airline passenger traffic using neural networks. *The Logistics and Transportation Review*, *31*(3), 239–252.

Nielsen. (2013). Under The Influence: Consumer Trust in Advertising. Retrieved from www.nielsen.com/us/en/insights/news/2013/under-the-influence-consumer-trust-in-advertising.html

Park, S., & Huh, S. (2019). A Social Network-Based Inference Model for Validating Customer Profile Data, *36*(4), 1217–1237.

Qiu, J., Lin, Z., & Li, Y. (2015). Predicting customer purchase behavior in the e-commerce context. *Electronic Commerce Research*, *15*(4), 427–452. https://doi.org/10.1007/s10660-015-9191-6

Reha Alpar. (2016). *Spor, Sağlık ve Eğitim Bilimlerinden Örneklerle Uygulamalı İstatistik ve Geçerlik - Güvenirlik*. Detay Yayıncılık. Retrieved from http://www.kitapyurdu.com/index.php?route=product/product&product_id=308595&gclid=EAIaIQobChMIzpPGzdet2QIVTrHtCh1CKQoOEAQYASABEgLTLfD_BwE

Sakar, C. O., Polat, S. O., Katircioglu, M., & Kastro, Y. (2018). Real-time prediction of online shoppers ' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. *Neural Computing and Applications*, *0*. https://doi.org/10.1007/s00521-018-3523-0

Song, Y. Y., & Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Archives of Psychiatry*, *27*(2), 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044

Ucar, M. K., & Topal, I. (2018). Rule-Based Determination of Chinese tourists to Turkey Opt Profile. In *2018 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1–4). IEEE. https://doi.org/10.1109/ASYU.2018.8553998

Vellido, A., Lisboa, P. J. G., & Meehan, K. (2015). Quantitative Characterization and Prediction of On-Line Purchasing Behavior : A Latent Variable Approach Approach, *4415*. https://doi.org/10.1080/10864415.2000.11518380

Wah, Y. B., Ismail, N. H., & Fong, S. (2011). Predicting car purchase intent using data mining approach. *Proceedings - 2011 8th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2011*, *3*, 1994–1999. https://doi.org/10.1109/FSKD.2011.6019863

Wong, B. K., Bodnovich, T. A., & Selvi, Y. (1996). Neural network applications in business: A review and analysis of the literature (1988-95), *19*, 301–320. https://doi.org/10.1016/S0167-9236(96)00070-X.