## IMISC 2018 Best Paper

# Evaluation of Call Center Efficiency Using Text Mining Approach

**Eda Üzüm[1]*, Seren Başaran[1], Sona Mardikyan[1]**

**1** Boğaziçi University, * Corresponding author, edauzumm@gmail.com

**Abstract**

In many business fields text mining methodologies are applied to analyze market, products, trends, quality, etc. Today, customer call center data are very valuable to understand customer needs and complaints, increase effectiveness and efficiency in technical services and customer loyalty, improve product quality and brand images. This study presents an application of text mining methods for customer call centers in a home appliance company. The dataset is provided by a home appliance company and includes 35 different country call center data. Random Forest and CART algorithms are applied to the recorded text which customers say directly to agents. According to the results, products' error causing parts are determined with a range of 49%-59% accuracy rate for different countries. The most used words prepared as a table and presented as a recommendation for a home appliance companies. As a result, this study identifies root cause of problems of call centers and how agents can deal with them to provide better services.

**Keywords:** Call center efficiency, Call center data, Text mining, Random forest, CART.

**Citation:** Üzüm, E., Başaran, S., Mardikyan, S. (2018, October) *Evaluation of Call Center Efficiency Using Text Mining Approach*. Paper presented at the Fifth International Management Information Systems Conference.

**Editor:** H. Kemal İlter, Ankara Yıldırım Beyazıt University, Turkey

**Received:** August 19, 2018, **Accepted:** October 18, 2018, **Published:** November 10, 2018



## Introduction

Today's business world is becoming more complex day by day and everything becomes data oriented. Making meaningful inferences from raw data is an important task. Undoubtfully, data mining is a fatal approach to make research and analysis to protect brand value. This study mainly focuses on creating insightful data and make a prediction with customers' complaints reported to the company's call centers. Call centers aims to solve customers' issues to increase customer satisfaction and loyalty by guiding during the call or directing technical services to customers. However, sometimes call center teams can make wrong guidance and this situation may lead to lose time, ineffective technical support, customer dissatisfaction and increment in cost of operations. In this study, prediction models were developed to evaluate the contents of call records by classification, using random forest and applying decision tree methods.

## Literature Review

Data mining is the computational process of discovering patterns in large datasets involving methods at the intersection of artificial intelligence, machine learning, statistics, and database systems (Kesavaraj & Sukumaran, 2013). The data is not only in form of figures, but also in text. Text mining is the automatic and semi-automatic extraction of implicit, previously unknown, and potentially useful information and patterns, from a large amount of unstructured textual data, such as natural-language texts (Chakraborty & Nagal, 2015). The main goal of text mining is to enable users to extract information from textual resources and deals with the operations like, retrieval, classification and summarization (Purohit, Atre, Jaswani, & Asawara, 2015; Gaikwad, Chaugule, & Patil, 2014; Khan, Husain, & Beg, 2015).

## Text Mining

Text summarization is used to reduce the length of a document while retaining most important points and general meaning. It is the process of collecting and producing concise representation of original text documents (Mukhedkar, Sakhare, & Kumar, 2016). Text summarization techniques can be applied on multiple documents at the same time. Quality and type of classifiers depend on nature and theme of the text documents (Al-Hashemi, 2010). Additional methods of text mining are introduced with standard text mining process to improve the relevance and accuracy of results (Chen & Zhang, 2014). Information extraction is an initial step for unstructured text analyzing (Nassif & Hruschka, 2013). Simplification of text is the work of information extraction.

In text mining, the typical text categorization process consists of pre-processing, indexing, dimensionally reduction, and classification (Lam, Ruiz, & Srinivasan, 1999; Lodhi, Saunders, Shawe-Taylor, Cristianini, & Watkins, 2002). The goal of categorization is to train classifier based on known examples and then unknown examples are categorized automatically. Information visualization techniques provide a convenient means to summarize documents in visual forms that allow users to fully understand and memorize data insights.

One of the study in the field of text mining was conducted by collecting real -life data from contact center summaries for a telecommunications company to learn various aspects of noise with detailed experiments to study its effect on automatic text classification systems (Agarwal, Godbole , Punjani, & Roy, 2007). El Gayar and Ghanem's (2012) study shows that general model that accepts any type of audio material and studies its content through machine learning techniques by automatically converting the audios into text and mining the text content. In Kaur & Singh's (2012) research, several text mining techniques on recorded telephone calls mimicking real agent/customer conversations after translating them into text in order to detect the speakers' emotions, and hence predict whether the customer is satisfied or dissatisfied of the service provided.

## Classification

On the other hand, classification is a data mining task of forecasting the value of a categorical variable by building a model based on one or more numerical and/or categorical variables (Kataria & Singh, 2013). There are various classification techniques like Decision Tree, Random Forest Classification, Bayesian Classification and Nearest Neighbour Classifier.

Decision trees are one of the most popular classification and prediction methods used in text mining. Decision tree recursively partitions the training dataset into smaller subdivisions based on a set of tests defined at each node or branch. Decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID) (Kaur & Singh, 2012; Kayri & Kayri, 2015).

Random Forest is a supervised learning algorithm. The "forest" it builds, is an ensemble of Decision Trees, most of the time trained with the "bagging" method. The general idea of the bagging method is that a combination of learning models increases the overall result (Donges, 2018). Random Forest is a classification model that tries to make more accurate classification by producing more compatible models using more than one decision tree. Each node is split using the best among a subset of predictors randomly chosen at that node (Liaw & Wiener, 2002). In addition, in random forests, there is no need for cross-validation or a separate test set to get an unbiased estimate of the test set error (Breiman & Cutler, 2002). It is estimated internally, during the run.

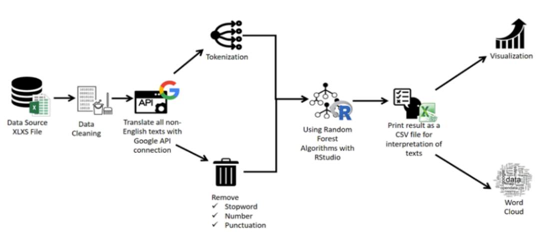


**Figure 1.** Architecture of the study.

## Data and Methodology

This study is prepared to make an improvement in the home appliance company's customer call centers to give more consistent respond to customers' complaints and needs by increasing prediction accuracy of error causing spare parts. All the steps including the data preprocessing is described with the Figure 1.

### Data

Data used in this study is provided by a company which is functioning in the home appliance industry. In data, there are records created by call center agents after making calls with customer, and shows how the technician estimates the related errors. Briefly, there are estimated error codes which are entered by the call center agents during the talk and the exact error codes which are recorded by the technician after the customer visit. The data collected between January 1, 2016 and April 31, 2017. The study is conducted by selecting a specific product which is dishwasher to limit the data, however the same analysis can be conducted to different products by selecting their related data. The dataset includes 35 different countries information which are Turkey, France, Hungary, Portugal, India, Belgium, Italy, Czech Republic, Turkmenistan, Kazakhstan, Ukraine, Greece, Romania, Netherlands, Singapore, Malaysia, Switzerland, Thailand, Hong Kong, Taiwan, Austria, Luxembourg, Peru, Chile, Germany, Bulgaria, Slovenia, Estonia, Croatia, Serbia, United Kingdom, Ireland, Australia, New Zealand, South Africa.

### Data Preprocessing

In data preprocessing step (Fig. 1), after reducing the length of document, information extraction is applied by transforming corpus of textual documents into more structured database, categorization is used for grouping words according to the spare parts and visualization is used for explaining data insights.

There are a lot of information about the call center in the data. However, there is no need to use every data available to achieve the desired result. For this reason, the parts that are not used in the project have been cleaned in the data to make the working model running faster. Moreover, records including incomplete fields are also removed. In final version of the dataset, the common characteristics related to customer's complaints about the company's dishwasher product line are included. These are customer's country, call duration, calling date, age of appliance, operating time, storage time, warranty, spare part, error causing part and reason of exchange or repair, information about summary of the customer's statements that has made in the phone conversation, and information about the part configuration where the spare part is caused by each error which is called "QM Part Structure". In addition to these fields, two main text columns used in this study are "Error Causing Part" and "Customer Text" which consists of sentences written in approximately 10 different language. These are Turkish, French, Italian, German, Bulgarian, Chinese, Spanish, Estonian, Croatian, Dutch. It is necessary to create a single databank to create inconsistencies in different language grouping. To eliminate language barriers, all sentences are translated into English through Google API, the corresponding code is written and then it is applied. The abbreviations, "-", "_" and similar shapes have been cleaned from the data. The length of data is reduced by removing whitespaces and punctuations for summarization through R and Excel tools. Punctuations also meaningless to understand the relationship between words and spare part. Removing punctuation will not only accelerate the data while running but also provide space. Changing words cases took an advantage to catch letters before translating words from many different languages to English. After the Google API is applied, stop words are removed to select more appropriate words. Stop words usually refers to the most common words in a language. In this dataset stop words in English language are eliminated. Also, numbers are eliminated because numbers will not show the correlation between customers' words during call and spare parts.

Tokenization is a method of separating clauses as words, expressions, symbols, and markers. The main purpose of this study is to find certain words in estimating defective spare part errors. For this reason, the clauses in the data are separated first by word. Each word forms the head of the columns. Later on, the lines below have been given in the form of 1-0. For example, there are two sentences;

- I will go home
- I will do my homework

The result of this sentences in terms of tokenization is:

| I | will | go | home | do | my | homework |
|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 1 | 1 | 0 | 0 | 1 | 1 | 1 |

In addition, the number of records that seldom lasted in the records in the data set is reduced. In this study the frequency level is set to 0.99. This can be expressed as; if there is a word having less than 1% of frequency in the data, it is removed from the data.

## Methodology

In this study, it is planned to utilize the prediction models in home appliance company customer call centers to estimate the correct failed spare parts to increase accuracy, which is the percentage of how close an estimated value is to the actual value. Also, it will identify root cause of problems by looking at controversial topics and how agents can deal with them to provide better solutions. In the study, text mining classification models are built with random forest and decision tree models after the data preprocessing step. Models try to predict the correct words with matching spare parts. All the steps are given in Fig. 1.

After the data preprocessing, random forest methodology is implemented to find which words refer which spare parts. All the translated data is represented in a matrix form as it is explained in tokenization part. Each record refers a different call and each column represents a distinct translated word. Next to the words a new column including the exact error causing part code is added. Thus, the necessary dataset is established for modeling. Data is partitioned for training and testing and random forest method is applied. The training partition is typically the largest partition, and contains the data used to build the various models we are examining. The same training partition is generally used to develop multiple models. Accuracy is calculated based on the proportions of the modeled and predicted values relative to the generated train and test data. The higher the degree of proximity to the ratio of 1, the higher the accuracy of the value established.

After the text mining approach, a decision tree is created between the variables that appear to be related to the error causing part field. These variables are age of appliance, operating time, storage time and country. Age of appliance refers to age of product and its unit is year. Operating time refers to time from when the product was sold to when the problem occurred and its unit is year. Storage time refers to when the product is stored until it is sold after and its unit is year. The most important factor in selecting these columns is the closest and meaningful columns in the data when creating a rule. With this algorithm, different rules can be determined for each country. In this study, two different decision trees are formed. One is based on the data of a selected country which is Turkey and the other is using all countries' data.

## Results

Visualization is very important to fully understand and memorize data insights word clouds are a technique that provides colorful visualization. Word cloud selects the frequent repetitions of the words in a script, bringing them together on a schematic level and making the most used ones bigger. In this study, the purpose of using it is to make use of the frequently used expressions of the words that the customers use. As shown in the Figure 2 and Figure 3 "water" word is used most commonly while customer describing own complaints to call center agents.

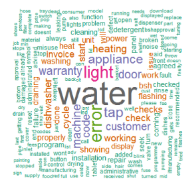


**Figure 2.** Word cloud of whole data.


**Figure 3.** Word cloud of Turkey.

The main purpose is to estimate spare parts errors of the dishwasher. Table 1. Illustrates some of the spare parts and their referring words derived from output of the Random Forest algorithm. The header includes name of spare parts which is written as bold, the rows include referring words. The whole output includes all the spare parts and their referring words.

**Table 1.** Words list from random forest.

| **Altern.water distrib** | **Aquastop** | **Cable harness** | **Crockery basket** | **Dispenser** | **Door-inner** |
|---|---|---|---|---|---|
| clean | power | circuit | bottom | aid | bottom |
| closed | trips | door | back | burned | damaged |
| detergent | broken | faulty | seal | basket | door |
| dish | defect | | side | broken | light |
| dishwasher | dishwasher | | tray | button | |
| door | fault | | | compartment | |
| exposed | function | | | cycle | |
| faulty | functioning | | | defective | |
| food | fuse | | | detergent | |
| indicates | hose | | | dispenser | |
| intervention | inlet | | | flap | |
| plug | inside | | | inside | |
| power | installation | | | installation | |
| washes | leaking | | | intervention | |
| water | light | | | leaking | |
| | noise | | | lid | |
| | panel | | | light | |
| | pipe | | | lock | |
| | plug | | | onoff | |
| | seal | | | programme | |
| | | | | releasing | |
| | | | | reset | |
| | | | | rinse | |
| | | | | rinsing | |
| | | | | shut | |
| | | | | soap | |
| | | | | stiff | |
| | | | | tablet | |
| | | | | tray | |

The accuracy rate for the words ranges from 49% to 59%. Since the test data is randomized every time the program runs, the accuracy rate for the words changes each time.

As a second approach to predict QM Part Structure categories which group spare parts that are related to each other CART algorithm is applied in R tool to create a decision tree graph for the user. The purpose of using it in this study is to get a meaningful decision tree by taking age of appliance, operating time, storage time and country as input variables. The rules derived from the CART algorithm for each spare part category.

For example;

```
Country: GB, IE, NZ
→ Operating Time (year): 1, 2
→ Age of Appliance (year): 11,12,13,14,15
→ QM Part Structure: PG_0210:04
```

Decision trees can be constructed by selecting only one country's data. Another Decision Tree for Turkey is generated to gain insights about the spare parts in Turkey.

For example;

```
Age of Appliance (year): 7, 8, 9,10,11,12
→ Storage Time (year): 5, 11, 12
→ Operating Time (year): 0,2,3,5
→ QM Part Structure: PG_0110:012
```

## Conclusions

The methodology is followed in this study is Text Mining process with Random Forest and Decision Tree techniques. In data cleaning and preprocessing step, empty records are cleaned, sloppy records which come from the Call Center Agents' manual entries are recovered. Moreover, stop words, punctuations, and numbers are removed in this step. Different text mining techniques are applied for creating an appropriate model. In modeling step, the aim is to find most accurate model that words refer error causing part of the products successfully. Random Forest and CART Decision Tree are applied. Since Random Forest is the best way to compare words as a 1-0 matrix according to target values. According to the Random Forest results, Call Center Agents will be able to predict which spare part has a problem with a faster and more efficient way. More accurate forecasting will increase customer satisfaction, and solving the problem at once will be less costly to the company. Additionally, CART Decision Tree helps to create a rule with all data in a more accurate way in terms of specific variables which are country, age of appliance, storage time and operating time. In summary, the customer text of the call center data has quite important to interpret which part has an error in the product. In addition, different countries have different problems about relevant products. The study helps companies to make more valuable interaction with customers.

## References


Agarwal, S., Godbole , S., Punjani, D., & Roy, S. (2007). How Much Noise Is Too Much: A Study in Automatic Text Classification. Seventh IEEE International Conference on Data Mining (pp. 3-12). Omaha: IEEE. doi:10.1109/ICDM.2007.21

Al-Hashemi, R. (2010). Text Summarization Extraction System (TSES). International Arab Journal of e-Technology, 1, 164-168.

Anuradha Purohit, D. A. (2015). Text Classification in Data Mining. Department of Computer Technology and Applications, Shri G.S. Institute of Technology and Science, Indore (M.P).

Breiman, L., & Cutler, A. (2002). Random Forests. Retrieved from University of California, Berkeley Department of Statistic: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm

Chakraborty, P., & Nagal, A. (2015). Software Innovations in Clinical Drug Development and Safety. Hershey PA.: IGI Global.

Chen, C., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques. Information Sciences, 314-347. doi:2014.01.015

Donges, N. (2018). The Random Forest Algorithm. Retrieved from Data Science Web Site: https://towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd

Friedl, M., & Brodley, C. (1997). Decision Tree Classification of Land Cover from Remotely Sensed Data. Remote Sensing of Environment, 399-409.

Gaikwad, S. V. (2014 ). Text Mining Methods and Techniques. International Journal of Computer Applications .

Kataria, A., & Singh, M. (2013). A Review of Data Classification Using K-Nearest Neighbour Algorithm. International Journal of Emerging Technology and Advanced Engineering, 3.

Kesavaraj, G., & Sukumaran, S. (2013). A study on classification techniques in data mining. 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT) (pp. 1-7). Tiruchengode: IEEE.

Lam, R. &. (1999). Automatic Text Categorization and Its Application to Text Retrieval. IEEE Transactions on Knowledge and Data Engineering .

Lam, W., Ruiz, M., & Srinivasan, P. (1999). Automatic Text Categorization and Its Application. IEEE Trans. Knowledge and Data Eng., 11, 865-879.

Lodhi, H., Saunders, C., Shawe-Taylor, J., Cristianini, N., & Watkins, C. (2002). Text Classification using String Kernels. Journal of Machine Learning Research 2, 419-444.

Lodhi, S. S.-T. (2002). Text Classification using String Kernels. Department of Computer Science, Royal Holloway, University of London.

Mukhedkar, B. A., Sakhare, D., & Kumar, R. (2016). Pragmatic Analysis Based Document Summarization. International Journal of Computer Science and Information Security (IJCSIS), 14, 145.

Nassif, L., & Hruschka, E. (2013). Document Clustering for Forensic Analysis: An Approach for Improving Computer Inspection. Information Forensics and Security IEEE Transactions, 8, 46-54.

Yogapreethi, N., & Maheswari, S. (2016, August). A Review on Text Mining in Data Mining Full Text. International Journal on Soft Computing (IJSC), 7.