Customer Churn Prediction Model and Identifying Features to Increase Customer Retention based on User Generated Content

Essam Abou el Kassem[1]
Mathematics and Statistics Department
Faculty of Commerce and Business Administration
Helwan University
Cairo, Egypt

Alaa Mostafa Abdelrahman[3]
Business Information Systems Department
Faculty of Commerce and Business Administration
Helwan University
Cairo, Egypt

Shereen Ali Hussein[2]
Management Department
Faculty of Commerce and Business Administration
Helwan University
Cairo, Egypt

Fahad Kamal Alsheref[4]
Information Systems Department
Faculty of Computers and Information
Beni-Suef University
Beni-Suef, Egypt

Abstract—**Customer churn is a problem for most companies because it affects the revenues of the company when a customer switch from a service provider company to another in the telecom sector. For solving this problem we put two main approaches: the first one is identifying the main factors that affect customers churn, the second one is detecting the customers that have a high probability to churn through analyzing social media. For the first approach we build a dataset through practical questionnaires and analyzing them by using machine learning algorithms like Deep Learning, Logistic Regression, and Naïve Bayes algorithms. The second approach is customer churn prediction model through analyzing their opinions through their user-generated content (UGC) like comments, posts, messages, and products or services' reviews. For analyzing the UGC we used Sentiment analysis for finding the text polarity (negative/positive). The results show that the used algorithms had the same accuracy but differ in arrangement of attributes according to their weights in the decision.**

Keywords—*Customer churn; telecom sector; churn prediction; sentiment analysis; machine learning; customer retention*

I. INTRODUCTION

During the enormous increase in numbers of customers who are using the communication sector and in numbers of companies [1], the competitive level between companies raised [2, 3]. Each company tries to survive in this competition through many strategies [4], The Main strategies are: 1) upsell existing customer, 2) increase duration of retention of their customers, 3) acquire new customers. Companies are concerned about seeking to keep or retain their customers as they are considered that as a profit, and it is cheaper to keep them than to earn a new one. Each company tries to keep its customers, by make them more loyal. Customers are great ambassadors in the market [5] as the company can use them for making advertising of the company's product or service.

This free advertising will cost nothing except the high quality and service after the sale.

Each company should be concerned with customer churn prediction and predict customers who are likely to leave the company to preserve its revenue, but companies must do it early [6].

Many types of research assured that machine learning is efficient in prediction through learning from past situations or previous data [7, 8, 9].

A customer churn happens when customers are not satisfied with a service provided by a company. It results in customers switching to another service provider. Customers have different reasons for churn, and all of them should not be treated in the same way. There is a need for a prediction model to predict churn customers and provide a strategy of retention depends on their churn factors [7].

According to vast numbers of people who use social media to show their opinions, whether by text, emotion, picture or video, we use sentiment analysis to analyze and classify every comment into positive, neutral, or negative [10]. And then, we track negative comments to the customer retention department to retain the churned customer.

In this paper, we proposed two approaches for helping companies to keep their customers by identifying the top reasons for churn and predicting the customer before the churn action is taken.

The rest of this paper is organized as: Section 2 presents the literature review. The background of the technique is presented in Section 3. The proposed model for churn detection from social media is briefly explained in Section 4. Experimental results and analysis are discussed in Section 5. We conclude the proposed study in Section 6.

II. LITERATURE REVIEW

This section presents the literature related to customers' retention and its prediction methods, identifying many factors related to customer retention, and using social media to get users' opinions to enhance retention.

Don Jyh-Fu Jeng, Thomas Bailey, has used hybrid, multiple criteria decision-making (MCDM) method to inspect customer retention framework and they found that the most common response was to look at pricing and customer service [11]. Ali Tamaddoni Jahromi, Stanislav Stakhovych, Michael Ewing used models for churn predictions in a B2B context, and to increase the profitability of retention campaigns; he found that boosting model, logistic regression, cost-sensitive and straightforward decision tree is applied on tests [12]. While Nitish Varshney and S.K. Gupta used social media analytics to get users' opinions through Twitter. The tweets were classified into three categories using a lexicon-based classifier and applying the association rule mining to find the dominant churn factor [13]. J. Vijaya, E. Sivasankar and S. Gayathri have proposed that ensemble classification techniques with hybrid fuzzy clustering provide more accuracy and better performance than single classifiers and clustering [14]. Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S used a distance factor in classifier decision [5]. Hossain, M.A, Chowdhury, M .R., & Jahan, N. have supported the importance of customer satisfaction in building a relationship between buyer and seller, using a model that was constrained only four constructs (price, network, customer care, and brand image) to explain customer satisfaction [15]. Adnan, Sajid, Awais, M.Nawaz, K.Alawfi, Amir, and Kaizhu have proposed a practical approach to classify, predict and extract important decision rules related to customer churn or not according to an intelligent rule-based decision-making technique, this technique based on rough set theory (RST). Experiments are carried out to evaluate the performance of RST using Exhaustive Algorithm (EA), Genetic Algorithm (GA), Covering Algorithm (CA), and the LEM2 algorithm (LA). Results show that RST based on GA is the most efficient technique for extracting implicit knowledge in the form of decision rules [16]. However, J. Vijaya and E. Sivasankar used RST with other techniques such as Bagging, Random Subspace, and Boosting. Boosting has achieved the highest accuracy of 93.73%. They found that ensemble classification techniques work better with a classification accuracy of 95.13% compared to any single model [17]. Despite Abhishek and Ratnesh have trained four machine learning models which are Logistic Regression, Support Vector Machine, Random Forest and Gradients boosted tree, and they found that Gradient boosted tree is best among other models, Both Random forest and Logistic regression are an average while SVM is underperforming between these models [18].

Most of the previous researches used classical machine learning techniques, and because of that, we tried to use the recent techniques like deep learning for identifying churn factors accurately. Also, we used the power of social medial for early churn detection of customers through using sentiment analysis that used in many areas of business analytics models.

In Section 3, we will discuss the algorithms we used in our study and sentiment analysis technique.

III. BACKGROUND OF TECHNIQUE

We have many algorithms in machine learning; we discuss only three algorithms that we used in our study and sentiment analysis technique.

A. Machine Learning Techniques-Classification Methods

Many approaches were applied to predict churn in the telecom sector; most of them have used machine learning technology and data mining. Techniques supposed for use in customer churn prediction:

a) Deep Learning Algorithm

Deep learning is a subset of machine learning based on neural networks that permit a machine to train itself to perform a task [19].

b) Naïve Bayes

Naïve Bayes classifier, also known as simple Bayes or independence Bayes, is a simple probabilistic classifier. This method builds on independence between the input variables, but it performs well even under conditions that might be considered suboptimal for algorithms [20, 21].

c) Logistic Regression

Logistic Regression is a traditional machine learning algorithm developed by a statistician. It is used for classification problems as it works through predictions of the relationship between the predictor variable and the output variable [21, 22].

B. Sentiment Analysis

Sentiment analysis, also known as opinion analysis, is one of the most important techniques used in social media. This technique is used to extract expressions, opinions of internet users, which are expressed in several forms (such as emotions, texts, pictures, videos), analyzing the opinion-oriented, and then classify it into positive, neutral, or negative sentiments [23, 24].

IV. RESEARCH CONTRIBUTION

A. Customer Churn Prediction Model

This paper proposed two main contributions; the first one is a model for customer Churn prediction by analyzing user-generated content, and the second model is identifying main attributes that help the retention department to keep their customers and prevent them from the churn.

Customer churn prediction model using UGC proposed in Fig. 1, the proposed model consists of multiple processes, as shown in Fig. 1; these steps are:

Step 1: User creates his user-generated content; this content could be post, opinion, or comments.

Step 2: English treebank applies text preprocessing, stemming, and lemmatization on English text to extract essential words in their basic form.
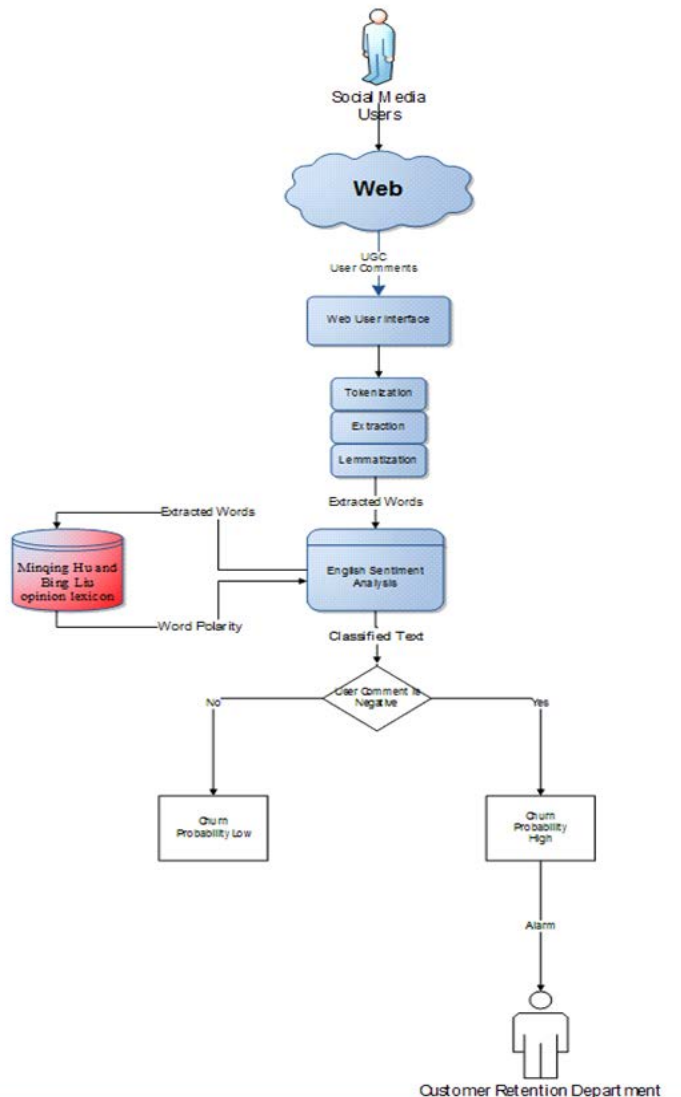
Fig 1.   Proposed Sentiment Analysis Model.

Step 3: Sentiment analysis classification: The extracted English words are entered for the classification process by measuring the polarity of each word then each text is classified according to its similarity with each class (positive, negative, neutral).

Step 4: Classify a user's comment, whether positive or negative; if the comment is adverse and churn probability is high, there is an alarm sent to the customer retention department with the user's id to communicate with the customer and try to retain the customer.

Step 5: The output of the proposed model is the sentimentally classified English text.

For example, I hate this company.

The model divides this sentence into four parts then classifies each word whether positive, negative, or neutral. Table I is shown an examples for sentence classification.

TABLE I.   EXAMPLE_SENTENCE_CLASSIFICATION

Example 1	I	hate	this	Company	Polarity Percentage
Classification	Neutral	Negative	Neutral	Neutral	Negative 99.8%
Example 2	Vodafone	is	a good	network	
Classification	Neutral	Neutral	Positive	Neutral	Positive 94%
Example 3	First	improve	your	Customer service	
Classification	Negative	Positive	Positive	Negative	Negative 92.7%

So, in the first example we have one-word negative and do not have any positive word, so the comment classified as a negative comment. While in the second example we have one-word positive and do not have any negative word, so the comment classified as a positive comment. Then in the last example we two positive and two negative words, the comment classified as negative. The classification happens for both the words and the sentence to identify the polarity to be more accurate.

B. Identifying Customer Churn Factors

We created a questionnaire and distributed it among customers for building a dataset that can be analyzed by machine learning algorithms. This dataset is built within different telecom companies and put almost known factors in it that later affect their decision, whether churn or not.

Description of survey's attributes

Gender: Gender of the customer.
Age: age of the customer.
The company of cellular communication service: name of the company which the customer-related.
For how many years a customer in the company: no. of years that customer relates to the company.
Customer's line type: line type, whether personal, business, or corporate.
Use to make a complaint: customers have used in displaying a complaint.
A problem happened with the company: mention a problem that happened.
The company tried to retain the customer: mention the response of the company.
Churn: customer ports out.

In the next section, we will present experimental results and analysis of proposed solutions.

V.   EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we have explored the experiments and results of the proposed study, our experiments are divided into two experiments. The first experiment analyzed users' comments from social media through sentiment analysis. The second experiment analyzed the dataset by using Naïve Bayes, Logistic Regression, and Deep learning. Then we used the correlation coefficient for finding the most effective factor with churn decision.

A. Experiments

Firstly, we targeted online social media users' comments from the Vodafone UK page. Then we collected many comments and used sentiment analysis to analyze and classify their comments according to positive, negative, or neutral sentiment. We collected 352 comments over six months, the comments from both males and females, most of them within 20-35 years.

Table II is shown a screenshot of the comments used in sentiment analysis.

TABLE II.   SOCIAL_MEDIA_COMMENTS

Post_ID	Comment_ID	Comment
1	1	I'd love to see how 3g works in my home
1	2	5G is not safe. So far nearly 250 EMF scientists from over 40 nations have signed this urgent appeal.
1	3	When is this coming to np13, south Wales?

Secondly, we have built our survey then distributed it among customers within four telecom companies. Each survey contains many questions in Fig. 2, such as:

Fig 2.   Screenshot for the Survey.

Then, we analyzed our survey through three algorithms (DL, NB, LR) to find the dominant churn factor.

B. Results of the Experiment

Here are the results of both customers' comments and dataset analysis.

a) Results of Comments

There are some comments from what we collected from social media (i.e. Vodafone UK Facebook page) and they are shown in Table III.

TABLE III.   COMMENT SENTIMENT EXAMPLE

Post_ID	User_ID	Comment	Proposed Model Classification
1	1	I'd love to see how 3g works in my home	Neutral 77.9%
1	2	5G is not safe. So far nearly 250 EMF scientists from over 40 nations have signed this urgent appeal.	Negative 51.8%
5	4	NEVER EVER use Vodafone. Liars, obstructive, and impossible to deal with.	Negative 99.7%
7	3	Vodafone is the best! 5G around the corner. Our best network provider □	Positive 96.0%
13	1	First, improve your customer service	Negative 92.7%
19	5	Vodafone is the best□❤□ network provider. Love it since 2002❤□□	Positive 98.4%
112	2	Great company. I've had faultless home broadband and mobile service from them.	Positive 95.4%

b) Result of Survey

And here is the result of the classified survey. Table IV is shown result by using NB algorithm.

TABLE IV.   NAIVE BAYES

Attribute	Weight
Your line type	0.152981
The company tried to retain you	0.080407
age	0.067216
A problem happened for you with the company	0.056266
Gender	0.040690
use to make a complaint	0.035092
For how many years are you a customer of the company	0.034172

Table V shows result by using LR algorithm.

TABLE V.   LOGISTIC REGRESSION

Attribute	Weight
Gender	0.090266
use to make a complaint	0.082144
The company tried to retain you	0.075134
For how many years are you a customer of the company	0.072974
A problem happened for you with the company	0.057011
age	0.041539
Your line type	0.005700

Table VI is shown result by using DL algorithm.

TABLE VI.   DEEP LEARNING

Attribute	Weight
age	0.1332652
Your line type	0.1158231
For how many years are you a customer of the company	0.0798538
use to make complaint	0.0781901
Gender	0.0721258
A problem happened for you with the company	0.0662030
The company tried to retain you	0.0394301

c) Result of Correlation

After we experimented with the algorithms, we used correlation for each algorithm in order to calculate the weight of attributes with the label attribute (churn decision whether yes or no) to know the value with the highest weight which affects customers' decision to be more accurate.

$$r = \frac{(X(i)-X\prime).(Y(i)-Y\prime)}{(n-1).S(X).S(Y)} \qquad (1)$$

Table VII shows a screenshot for correlation Naive Bayes table, the rest of the table in Appendix B.

TABLE VII.   COR_NAIVE_BAYES

Attribute	Weight
A problem happened for you with the company = Withdraw from the credit without using it	0.1673900
The company tried to retain you = Offered minutes for free	0.1556922
A problem happened for you with the company = The internet was cut	0.1321921
The company tried to retain you = Offered more internet megabits	0.1321921
A problem happened for you with the company = Problem in payment plan	0.1096181
A problem happened for you with the company = The internet was very slowly	0.1096181
The company tried to retain you = Offered free gigabits	0.1096181
The company tried to retain you = Reduce the invoice	0.1096181
The company tried to retain you = By offers	0.1068224
age = 36-45	0.1027885

TABLE VIII.   COR_DEEP_LEARNING

Attribute	Weight
A problem happened for you with the company = Withdraw from the credit without using it	0.167390
The company tried to retain you = Offered minutes for free	0.155692
A problem happened for you with the company = The internet was cut	0.132192
The company tried to retain you = Offered more internet megabits	0.132192
A problem happened for you with the company = Problem in payment plan	0.109618
A problem happened for you with the company = The internet was very slowly	0.109618
The company tried to retain you = Offered free gigabits	0.109618
The company tried to retain you = Reduce the invoice	0.109618
The company tried to retain you = By offers	0.106822
age = 36-45	0.102789

TABLE IX.   COR_LOGISTIC_REGRESSION

Attribute	Weight
A problem happened for you with the company = Withdraw from the credit without using it	0.16739
The company tried to retain you = Offered minutes for free	0.15569
A problem happened for you with the company = The internet was cut	0.13219
The company tried to retain you = Offered more internet megabits	0.13219
A problem happened for you with the company = Problem in payment plan	0.10962
A problem happened for you with the company = The internet was very slowly	0.10962
The company tried to retain you = Offered free gigabits	0.10962
The company tried to retain you = Reduce the invoice	0.10962
The company tried to retain you = By offers	0.10682
age = 36-45	0.10279

In Table VIII, is a screenshot for correlation Deep Learning table, the rest of the table is in Appendix C.

Screenshot for the correlation Logistic Regression table is shown in Table IX, the rest of the table is in Appendix D.

C. Experiment Results Summary

Three main points summarize our results:

a) UGC Analysis

The result showed the sentiment of each comment; it divided each sentence into parts then classified the kind of sentence whether positive, negative, or neutral with percentage.

b) Dataset Analysis

Here we can notice that the weights of attributes have differed from one algorithm to another. In Naïve Bayes, the line type of users has the highest weight while in Logistic Regression, gender has the highest weight, and in Deep Learning, the customer's age has the highest weight.

Whereas the years the customer has joined, the company has the lowest weight in Naïve Bayes. In contrast, line type of the customer has the lowest weight in the Logistic Regression, but in Deep Learning the trial of the company to retain its customer has the lowest weight.

c) Correlation Analysis

The results of correlation shown the values affect each attribute, which has a high rate on the customer's decision, whether churn or not churns from the company. We can notice that the most effective value in 'A problem happened for you with the company' attribute was 'Withdraw from the credit without using it' in Correlation Naïve Bayes, Correlation Deep Learning, and Correlation Logistic Regression, which means it has the highest weight in customers' decision. And after it came the best way to retain customers was offered minutes for free. And as we see all algorithms have the same order of values.

VI. CONCLUSION

The importance of this research paper comes from the importance of customer churn in the telecom sector. It helps companies to make more profit as customer churn is considered as one of the most important sources of income for the telecom sector. Hence, we build a model to analyze the behavior of customers and predict whom customers want to churn. In this study, we used Deep Learning, Naïve Bayes, and Logistic Regression algorithms for identifying the most important factors that affect the churning process. And we found the correlation for these algorithms to get the value of attribute which has the highest weight in the decision. Moreover, we analyzed the UGC by using sentiment analysis to analyze and classify customers' opinions. According to the survey and comments, we can track the churned customer to the customer retention department to retain him with the customer's id.

REFERENCES

[1] DOAA, M. E.; MOHAMED, H. A survey on sentiment analysis challenges. Journal of King Saud University-Engineering Sciences, 2016, 4.

[2] Gerpott TJ, Rams W, Schindler A. Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market. Telecommun Policy. 2001; 25:249—69.

[3] AHMAD, Abdelrahim Kasem; JAFAR, Assef; ALJOUMAA, Kadan. Customer churn prediction in telecom using machine learning in big data platform. Journal of Big Data, 2019, 6.1: 28.

[4] Wei CP, Chiu IT. Turning telecommunications call details to churn prediction: a data mining approach. Expert Syst Appl. 2002; 23(2):103—12.

[5] AMIN, Adnan, et al. Customer churn prediction in telecommunication industry using data certainty. Journal of Business Research, 2019, 94: 290-301.

[6] Qureshii SA, Rehman AS, Qamar AM, Kamal A, Rehman A. Telecommunication subscribers churn prediction model using machine learning. In: Eighth international conference on digital information management. 2013. P.131—6.

[7] ULLAH, Irfan, et al. A churn prediction model using random forest: analysis of machine learning techniques for churn prediction and factor identification in telecom sector. IEEE Access, 2019, 7: 60134-60149.

[8] Umayaparvathi V, Iyakutti K. A survey on customer churn prediction in telecom industry: datasets, methods and metric. Int Res J Eng Technol. 2016; 3(4):1065—70.

[9] Yu W, Jutla DN, Sivakumar SC. A churn-strategy alignment model for managers in mobile telecom. In: Communication networks and services research conference, vol. 3. 2005. p. 48—53.

[10] SALEENA, Nabizath, et al. An ensemble classification system for twitter sentiment analysis. Procedia computer science, 2018, 132: 937-946.

[11] SORIANO, Domingo Ribeiro; JENG, Don Jyh-Fu; BAILEY, Thomas. Assessing customer retention strategies in mobile telecommunications. Management Decision, 2012.

[12] JAHROMI, Ali Tamaddoni; STAKHOVYCH, Stanislav; EWING, Michael. Managing B2B customer churn, retention and profitability. Industrial Marketing Management, 2014, 43.7: 1258-1268.

[13] VARSHNEY, Nitish; GUPTA, S. K. Mining Churning Factors in Indian Telecommunication Sector Using Social Media Analytics. In: International Conference on Data Warehousing and Knowledge Discovery. Springer, Cham, 2014. p. 405-413.

[14] VIJAYA, J.; SIVASANKAR, E.; GAYATHRI, S. Fuzzy Clustering with Ensemble Classification Techniques to Improve the Customer Churn Prediction in Telecommunication Sector. In: Recent Developments in Machine Learning and Data Analytics. Springer, Singapore, 2019. p. 261-274.

[15] HOSSAIN, Md Alamgir, et al. Customer Retention and Telecommunications Services in Bangladesh. International Journal of Asian Social Science, 2017, 7.11: 921-930.

[16] AMIN, Adnan, et al. Customer churn prediction in the telecommunication sector using a rough set approach. Neurocomputing, 2017, 237: 242-254.

[17] VIJAYA, J.; SIVASANKAR, E. Computing efficient features using rough set theory combined with ensemble classification techniques to improve the customer churn prediction in telecommunication sector. Computing, 2018, 100.8: 839-860.

[18] GAUR, Abhishek; DUBEY, Ratnesh. Predicting Customer Churn Prediction In Telecom Sector Using Various Machine Learning Techniques. In: 2018 International Conference on Advanced Computation and Telecommunication (ICACAT). IEEE, 2018. p. 1-5.

[19] Abdalla, Rifaat, and Marwa Esmail. WebGIS for Disaster Management and Emergency Response. Springer, 2018. p. 59.

[20] VAN DER HEIDE, E. M. M., et al. Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. Journal of dairy science, 2019, 102.10: 9409-9421.

[21] VAFEIADIS, Thanasis, et al. A comparison of machine learning techniques for customer churn prediction. Simulation Modelling Practice and Theory, 2015, 55: 1-9.

[22] AL IMRAN, Abdullah; AMIN, Md Nur; JOHORA, Fatema Tuj. Classification of Chronic Kidney Disease using Logistic Regression, Feedforward Neural Network and Wide & Deep Learning. In: 2018 International Conference on Innovation in Engineering and Technology (ICIET). IEEE, 2018. p. 1-6.

[23] ALHARBI, Ahmed Sulaiman M.; DE DONCKER, Elise. Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information. Cognitive Systems Research, 2019, 54: 50-61.

[24] KOWCIKA, A., et al. Sentiment analysis for social media. International journal of advanced research in computer science and software engineering, 2013.

APPENDIX A

Gender *

○ Female

○ Male

age *

○ 15-20

○ 21-25

○ 26-35

○ 36-45

○ Above 45

The company of your cellular communication service *

○ Orange

○ Vodafone

○ Etisalat

○ We

For how many years are you a customer of the company *

○ Less than one year

○ 1-5

○ 5-10

○ More than 10

Your line type *

○ Personal

○ Business

○ Corporate

you use to make your complaint *

☐ Social Media

☐ Customer Service Representative

☐ Call Center Representative

A problem happened for you with the company *

○ The package was ended without inform me

○ Withdraw from the credit without using it

○ Rooming issues

○ The internet was very slowly

○ The internet was cut

○ Problem in mobile data

○ Stopped the service of calls or cellular data

○ Bad network

○ Cash service, the money late to send

○ cut the network

○ Way of payment for internet service

○ Extra money charged

○ Problem in the Sim card

○ Problem in recharging

○ Errors in the invoice

○ Anghami bundle or call tone without my acceptance

○ No problem happened

○ Problem in payment plan

The company tried to retain you *

○ Offered more internet megabits

○ Return back the credit

○ Offered a free bundle

○ Offered minutes for free

○ I didn't mention the problem

○ Nothing happened

○ Justified the reason and apologized

○ Regular maintenance

○ Offered an internet modem for free

○ Reduce the invoice

○ Changed the Sim card from 3G to 4G

○ Went to another store

○ Offered to pay only half of my bill for 6 month

○ Not enough time

○ Corrected the invoice

○ Brought my money back

○ Offered a better plan of payment

○ Made a discount on the new package and return the amount

○ By offers

○ Come back my flex

○ Gave me new Sim card

○ Solve the problem

○ Offered free gigabits

churn (port out) *

○ Yes

○ No

APPENDIX B

Attribute	Weight
A problem happened for you with the company = Withdraw from the credit without using it	0.1673900
The company tried to retain you = Offered minutes for free	0.1556922
A problem happened for you with the company = The internet was cut	0.1321921
The company tried to retain you = Offered more internet megabits	0.1321921
A problem happened for you with the company = Problem in payment plan	0.1096181
A problem happened for you with the company = The internet was very slowly	0.1096181
The company tried to retain you = Offered free gigabits	0.1096181
The company tried to retain you = Reduce the invoice	0.1096181
The company tried to retain you = By offers	0.1068224
age = 36-45	0.1027885
The company tried to retain you = Solve the problem	0.0924705
The company tried to retain you = Return back the credit	0.0850286
A problem happened for you with the company = Anghami bundle or call tone without my acceptance	0.0831868
For how many years are you a customer of the company = More than 10	0.0792394
age = 21-25	0.0790395
A problem happened for you with the company = Bad	0.0768866

Attribute	Weight
network	
age = Above 45	0.0761141
A problem happened for you with the company = Rooming issues	0.0740389
A problem happened for you with the company = Stopped the service of calls or cellular data	0.0740389
For how many years are you a customer of the company = Less than 5 years	0.0692804
use to make complaint = Call Center Representative	0.0676756
age = 26-35	0.0650416
A problem happened for you with the company = Problem in recharging	0.0638076
The company tried to retain you = Justified the reason and apologized	0.0638076
The company tried to retain you = Offered a free bundle	0.0638076
The company tried to retain you = Offered to pay only half of my bill for 6 month	0.0638076
A problem happened for you with the company = Errors in the invoice	0.0524717
A problem happened for you with the company = Extra money charged	0.0518476
A problem happened for you with the company = Problem in the Sim card	0.0518476
A problem happened for you with the company = The package was ended without inform me	0.0518476
The company tried to retain you = Brought my money back	0.0518476
The company tried to retain you = Gave me new Sim card	0.0518476
The company tried to retain you = Offered a better plan of payment	0.0518476
The company tried to retain you = Offered an internet modem for free	0.0518476
A problem happened for you with the company = Way of payment for internet service	0.0364868
The company tried to retain you = Come back my flex	0.0364868
The company tried to retain you = Corrected the invoice	0.0364868
The company tried to retain you = I didn't mention the problem	0.0364868
The company tried to retain you = Made a discount on the new package and return the amount	0.0364868
The company tried to retain you = Not enough time	0.0364868
The company tried to retain you = Regular maintenance	0.0364868
The company tried to retain you = Went to another store	0.0364868
Your line type = Business	0.0328753
use to make complaint = Customer Service Representative	0.0279953

The company tried to retain you = Nothing happened	0.0125065
Your line type = Personal	0.0107071
A problem happened for you with the company = Problem in mobile data	0.0107071
For how many years are you a customer of the company = From 5 to 10 years	0.0104629
Gender = Female	0.0065408

APPENDIX C

Attribute	Weight
A problem happened for you with the company = Withdraw from the credit without using it	0.167390
The company tried to retain you = Offered minutes for free	0.155692
A problem happened for you with the company = The internet was cut	0.132192
The company tried to retain you = Offered more internet megabits	0.132192
A problem happened for you with the company = Problem in payment plan	0.109618
A problem happened for you with the company = The internet was very slowly	0.109618
The company tried to retain you = Offered free gigabits	0.109618
The company tried to retain you = Reduce the invoice	0.109618
The company tried to retain you = By offers	0.106822
age = 36-45	0.102789
The company tried to retain you = Solve the problem	0.092471
The company tried to retain you = Return back the credit	0.085029
A problem happened for you with the company = Anghami bundle or call tone without my acceptance	0.083187
For how many years are you a customer of the company = More than 10	0.079239
age = 21-25	0.079040
A problem happened for you with the company = Bad network	0.076887
age = Above 45	0.076114
A problem happened for you with the company = Rooming issues	0.074039
A problem happened for you with the company = Stopped the service of calls or cellular data	0.074039
For how many years are you a customer of the company = Less than 5 years	0.069280
use to make complaint = Call Center Representative	0.067676
age = 26-35	0.065042

A problem happened for you with the company = Problem in recharging	0.063808
The company tried to retain you = Justified the reason and apologized	0.063808
The company tried to retain you = Offered a free bundle	0.063808
The company tried to retain you = Offered to pay only half of my bill for 6 month	0.063808
A problem happened for you with the company = Errors in the invoice	0.052472
A problem happened for you with the company = Extra money charged	0.051848
A problem happened for you with the company = Problem in the Sim card	0.051848
A problem happened for you with the company = The package was ended without inform me	0.051848
The company tried to retain you = Brought my money back	0.051848
The company tried to retain you = Gave me new Sim card	0.051848
The company tried to retain you = Offered a better plan of payment	0.051848
The company tried to retain you = Offered an internet modem for free	0.051848
A problem happened for you with the company = Way of payment for internet service	0.036487
The company tried to retain you = Come back my flex	0.036487
The company tried to retain you = Corrected the invoice	0.036487
The company tried to retain you = I didn't mention the problem	0.036487
The company tried to retain you = Made a discount on the new package and return the amount	0.036487
The company tried to retain you = Not enough time	0.036487
The company tried to retain you = Regular maintenance	0.036487
The company tried to retain you = Went to another store	0.036487
Your line type = Business	0.032875
use to make complaint = Customer Service Representative	0.027995
The company tried to retain you = Nothing happened	0.012507
Your line type = Personal	0.010707
A problem happened for you with the company = Problem in mobile data	0.010707
For how many years are you a customer of the company = From 5 to 10 years	0.010463
Gender = Female	0.006541

APPENDIX D

Attribute	Weight
A problem happened for you with the company = Withdraw from the credit without using it	0.16739
The company tried to retain you = Offered minutes for free	0.15569
A problem happened for you with the company = The internet was cut	0.13219
The company tried to retain you = Offered more internet megabits	0.13219
A problem happened for you with the company = Problem in payment plan	0.10962
A problem happened for you with the company = The internet was very slowly	0.10962
The company tried to retain you = Offered free gigabits	0.10962
The company tried to retain you = Reduce the invoice	0.10962
The company tried to retain you = By offers	0.10682
age = 36-45	0.10279
The company tried to retain you = Solve the problem	0.09247
The company tried to retain you = Return back the credit	0.08503
A problem happened for you with the company = Anghami bundle or call tone without my acceptance	0.08319
For how many years are you a customer of the company = More than 10	0.07924
age = 21-25	0.07904
A problem happened for you with the company = Bad network	0.07689
age = Above 45	0.07611
A problem happened for you with the company = Rooming issues	0.07404
A problem happened for you with the company = Stopped the service of calls or cellular data	0.07404
For how many years are you a customer of the company = Less than 5 years	0.06928
use to make complaint = Call Center Representative	0.06768
age = 26-35	0.06504
A problem happened for you with the company = Problem in recharging	0.06381
The company tried to retain you = Justified the reason and apologized	0.06381

Attribute	Weight
The company tried to retain you = Offered a free bundle	0.06381
The company tried to retain you = Offered to pay only half of my bill for 6 month	0.06381
A problem happened for you with the company = Errors in the invoice	0.05247
A problem happened for you with the company = Extra money charged	0.05185
A problem happened for you with the company = Problem in the Sim card	0.05185
A problem happened for you with the company = The package was ended without inform me	0.05185
The company tried to retain you = Brought my money back	0.05185
The company tried to retain you = Gave me new Sim card	0.05185
The company tried to retain you = Offered a better plan of payment	0.05185
The company tried to retain you = Offered an internet modem for free	0.05185
A problem happened for you with the company = Way of payment for internet service	0.03649
The company tried to retain you = Come back my flex	0.03649
The company tried to retain you = Corrected the invoice	0.03649
The company tried to retain you = I didn't mention the problem	0.03649
The company tried to retain you = Made a discount on the new package and return the amount	0.03649
The company tried to retain you = Not enough time	0.03649
The company tried to retain you = Regular maintenance	0.03649
The company tried to retain you = Went to another store	0.03649
Your line type = Business	0.03288
use to make complaint = Customer Service Representative	0.02800
The company tried to retain you = Nothing happened	0.01251
Your line type = Personal	0.01071
A problem happened for you with the company = Problem in mobile data	0.01071
For how many years are you a customer of the company = From 5 to 10 years	0.01046
Gender = Female	0.00654