Predicting Customer Behavior in Support Channels Using Machine Learning

Muhamed Begović, Elma Avdagić-Golub, Belma Memić and Amel Kosovac
University of Sarajevo/ Faculty of Traffic and Communication, Sarajevo, Bosnia and Herzegovina
muhamed.begovic@fsk.unsa.ba, elma.avdagic@fsk.unsa.ba, belma.memic@fsk.unsa.ba, amel.kosovac@fsk.unsa.ba

***Abstract* - Support channels represent a unique opportunity to improve customer satisfaction by offering a consistent experience in resolving customer issues. Several surveys show that customers have raised their standards of customer support services. While only a few years ago customers willingly waited a long time to speak with one of the service agents and were patient for their problem to be resolved, today's customers have very limited patience and want a solution to the problem immediately. Customers don't want to settle for a mediocre support channel experience. Support channels must provide superior service capacities so that customers see that the company values their choice and time. Efficient management of support centers implies accurate modeling of customer behavior on hold. The subject of our research is the application of data research techniques for predicting customer behavior in support channels. In this paper, we apply machine learning methods to predict customer behavior. Based on historical data in the service system, we use classification algorithms to predict customer patience in service channels.**

Keywords - support channels, machine learning, customer patience

I. Introduction

Service companies face huge competition in the global market. Strategic planning of long-term relations with customers is the key to business success. Contact centers are complex systems where customer requests arrive through different media, which we call support channels. Support channels are calls, e-mails, chatbots, live chats, etc. They have a direct impact on customer satisfaction with the company's service. According to the type of communication, we distinguish two types of support call channels: incoming and outgoing calls[1]. Incoming call channels are most often used to provide information to customers, aiming to solve their problems, while outgoing calls have the purpose of proactive communication with existing or potential customers[2]. Support channels are service systems. The support call channel service process consists of three basic components: the process of the customer's call arriving in the system, the service process, and the customer's waiting time[3]. The standard assumption is that call arrivals follow a Poisson process with a (potential) random arrival rate, assumed to be constant over a given time [4]. A service process is the duration of a conversation between a customer and a customer service representative (CSR). Waiting time is the time required for the customer to receive the necessary information. This time is called the time of patience. Patience time is defined as the time the customer is willing to wait for the service [3]. It depends on the relationship between demand and supply, that is, the number of incoming calls in relation to the service capacity of CSR. When calling, the customer first encounters the Interactive Voice Response (IVR) and triages for the requested services, providing input information. IVR systems for individual customers can be complex and a labyrinth from which the customer leaves the service system. Therefore, efficient management of the service system implies accurate modeling of customer behavior in the service system

Data science is one of the main drivers of changes in business processes, from the way we treat customers to modeling their behavior. Accurate predictive models for customer behavior are key to the design and optimization of business processes[5]. Predicting the customer's patience, i.e. predicting the waiting time of the upcoming call, incoming calls can be treated and routed automatically and intelligently. The biggest challenge of contact center management is to respond to each incoming customer call uniquely in real-time, taking care that all customers do not react in the same way while waiting in the service queue.

In this paper, we observe a telephone support channel with incoming calls. Each call consists of a series of attributes that describe the customer's interaction with the support channel, such as the date, time of the call, reason for the call, and others. We use the data of the user's interaction with the service system, such as the date and time of the call, and use them to describe previous contacts. This creates the possibility for profiling and segmenting customers according to this attribute in the support channel. The main goal of our work is to identify customer behavior patterns. Based on these patterns of behavior, we conclude the premature abandonment of the customer from the service system by applying machine learning. We predict customer behavior at the individual level, based on attributes that describe the customer's recent contacts, as well as patience to speak with CSR. This kind of work can provide instructions for practical application in all companies, especially service companies, through any support channel.

The rest of the paper is organized as follows: Section II presents existing research efforts in this study field, Section III presents the simulation model, and Section IV presents the results of the prediction model and the discussion. Finally, conclusions and guidelines for future research are presented.

II. LITERATURE REVIEW

Quantitative models for designing and managing service systems can be observed through three common analytical techniques: queuing theories, text mining models, and data mining[6]. Problems investigated in the mathematical modeling of call centers can be found in papers [7], [8], [9], and [10]. In particular, papers related to the application of the impatient waiting model in the management of service systems are [11], [12], and [13]. Text mining models refer to obtaining information from customers' text messages, such as detecting emotions, monitoring conversations between customers and CSRs to reject problematic calls, etc.[14], [15]. Data mining models are aimed to predict the intensity of call arrivals [16], [17], predicting the reason for the call based on previous interactions [18] and routing calls to the most appropriate CSR [2].

Customer patience is an attribute that describes behavior in a service system. Patience time is defined as the time the customer is willing to wait, and waiting time is defined as the time the customer needs to wait to receive the service [3]. Patience time is mostly interrupted, as the vast majority of customers get service before their patience time runs out. Among the factors that influence patience is the time already spent on the call, i.e. the total time spent in the IVR and the waiting time in the queue [19]. Positive waiting time is a manifestation of the mismatch between demand (customer service requests) and supply (availability of CSR). In the last case, such a wait results in system failure, i.e. customer abandonment [4].

Our model differs from the mentioned studies. First, the model focuses on predicting the behavior of customers in the queue. Second, the model includes attributes that characterize the customer's previous contacts through the telephone channel. Customers behave differently in the queue, and not all of them are capable of navigating the IVR system. Individual customers may have a greater tendency to opt out of the queue. Based on previous interactions, it can be assessed whether the customer tends to abandon the service system. To the best of our knowledge, our model is the first attempt to incorporate opt-out information from a customer's previous contacts, to predict their behavior in a future telephone channel approach.

III. METHODS

A. Dataset

For this research, we used the operational data of the bank's contact center. The data covers twelve months of one year for each phone call. The database consists of 444,448 phone calls, with a complete history of call movement through the contact center. In the dataset, we distinguish attributes that describe the current call (transaction), such as the date and time of the call, the reason for the call, and the time spent in the queue. These attributes are very important because they affect the customer's patience. As we can see in Fig. 1, the intensity of calls is higher at the beginning of the month, especially in the first week of the month.

In addition, it can be concluded that in the first and last week of the month, there is an increase in unanswered calls compared to answered ones. This means that maybe customers are more impatient or the CSR shift schedule is bad. Also, the system is the busiest around 9 a.m., then the number of incoming calls gradually decreases until around noon and increases again around 4 p.m., as shown in Fig. 2.

The reason for the customer's call is one of the crucial attributes for analyzing the customer's patience. We will consider the reason for the call as the type of problem reported by the customer. Each type has a different intensity of arrival and serving time, and thus the waiting time for a given service [20] [8], [21], [22], as shown in Fig. 3.

We used only those calls that are recognized in the system and have assigned Customer ID attributes since the goal of modeling customer behavior is to predict the class of patience based on customer characteristics.

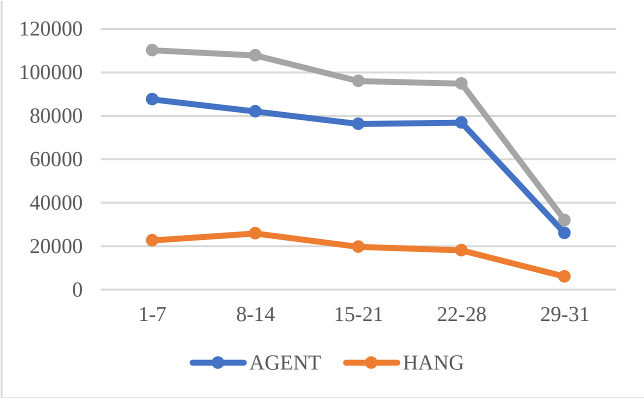

Figure 1. Call volume for one month

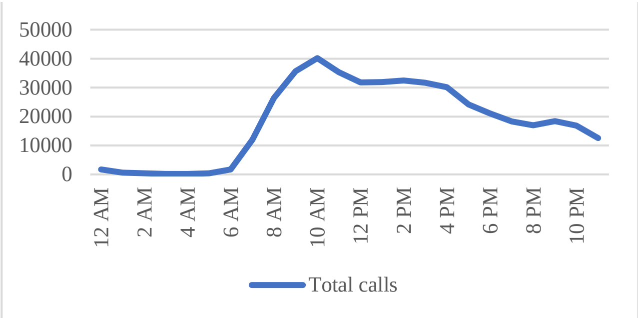

Figure 2. Call volume per hour

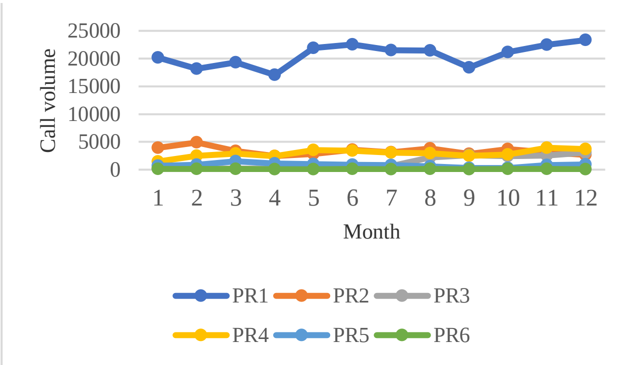

Figure 3. Call volume per type of problem

Our dataset contains 15 columns, and 192,889 individual records (customer calls). 154,441 calls have an "agent" call outcome, and 38,447 calls have a "hang" outcome. The first 6 columns represent categorical variables and need to be encoded. The categorical variables were encoded to convert them into a numerical format so that they can be used as input features in the logistic regression model.

We used One-Hot Encoding, where each unique value in a categorical column is represented as a new column and is assigned a binary value of either 1 or 0, depending on its presence in the row. Each row would have values of either 1 or 0 for these columns, indicating the presence or absence of the respective value in the row. The remaining columns are numerical. The final column, is the target variable that needs to be predicted and it has two possible values: "agent" and "hang". The representation of individual classes in the target variable was uneven in the ratio of 8:1 in favor of the "agent" class. Therefore, we equalized the number of samples for classes and the resulting dataset had 76,894 records.

The data was split into a training set and a testing set, with 80% of the data being used for training and the remaining 20% for testing. The model was fine-tuned using grid search to determine the best hyperparameters for the solver, penalty, and regularization strength.

B. *Prediction of the customer's patience using classification algorithms*

The paper aims to develop a predictive model based on customer characteristics to predict the customer's patience in the service queue during the next call to the contact center, including previous interactions. The target variable for this algorithm is the customer's patience class. So, our problem is to predict whether the call belongs to call class 1 (if the customer will talk to CSR) or 0 (if the customer will leave the service system - hang up).

Let's denote the total number of calls with I, for each call $i \in \{1,2,...I\}$ we define the variable y_i as:

$$y_i|x_i = \begin{cases} 1 & \text{if a customer was served by the agent} \\ 0 & \text{if the customer abandoned the system} \end{cases}, \quad (1)$$

where x_i denotes a vector of customer attributes. Therefore, we divide customers into patient and impatient ones. To build a predictive model for the target variable, we used logistic regression, a supervised machine learning algorithm commonly used for binary classification problems.

Referring to linear regression and model training, class probability y_i for characteristic x_i is:

$$p(y_i|x_i) = w_0x_0 + w_1x_1 + \cdots w_I x_I\,, \quad (2)$$

where w is the weight of each feature per class. The vector of characteristics represents the profile of the customer. Customer patience is the aggregate result of assessments based on fourteen characteristics. The attributes used in our analysis are shown in Table I. Customer profile attributes are divided into categorical characteristics and numerical values. Each categorical attribute in the data is modeled by a binary characteristic.

Some attributes require special modeling, which was done in the paper [23]. The attributes of the customer's job, marital status, education, and contact are static and represent the customer's personal characteristics. The attribute of customer priority indicates the importance of the customer to the company.

Customers have different preferences, expectations, and needs from the company, but they also have different incomes and spending profiles, so they must be managed in different ways [24]. This is especially evident in situations when the company changes its business strategies or solves management problems. Not all customers will react the same. This attribute has three categories, normal, priority, and higher priority.

We include four attributes to represent the date and time of past contacts, namely, the month, the regular week of the call, the day period of the call, and the elapsed time since the last call measured in days. The ordinal number of the week shows in which working week in a given month the call is made. Following the example of [25] the day is divided into relatively short time blocks, short enough that the call arrival rate does not change significantly in the block. The day is divided into five segments (time blocks): 7:00 - 11:00, 11:00 - 14:00, 14:00-17:00, 17:00-20:00, and 20:00-24:00.

We also distinguish attributes that show the frequency of recent contacts with the company, i.e. the frequency of calls within a week and the frequency of calls within the same day. The customer's patience can be affected by the period during which the problem is unresolved and how often the customer has to call to get it resolved.

The type of Call is the reason why the customer calls customer support. This is the type of service that the customer requests from the company. The reason for the call can be a decisive factor in the customer's patience. If the problem is urgent to solve, the customer will be ready to wait even longer [19].

The last attribute in the observation is the waiting time. It is important to distinguish between the time the customer has to wait before contacting the CSR and the time the customer is willing to wait before leaving the system. The first time is set by the system according to the number of CSR that are available. The time the customer is willing to wait is a measure of the customer's patience. Clearly, the time before being served is not known in advance. Therefore, in order to be able to use this time as an input variable, it is necessary to make an estimate earlier. The assessment of that time was not the subject of our work. However, this can be made based on parameters such as the intensity of incoming calls, the number of available agents, the period of the day, etc. In addition, it is possible to use advanced methods and machine learning such as those in the papers [26], [27].

We especially emphasize that the target variable (1) is defined individually for each call. Based on previous calls, it is necessary to conclude the characteristics of the customer's patience for the current call. The steps of the prediction algorithm implementation are presented in Fig. 4.

For the proposed model, the algorithm code was written with the help of the Google Collaboratory platform as part of the Google Research project. It is an online platform that enables the implementation of Python code. The Google Collaboratory platform enables the implementation of machine learning algorithms with the use of the necessary processing and memory capacities, which, in the case of large databases, are very difficult to provide on standard computers.

TABLE I. CUSTOMER PROFILE ATTRIBUTES

Customer profile		
Attribute	Data type	Sample
Customer ID	Integer	25365864
Customer job	String /Category	management
Customer marital status	String /Category	single
Customer education	String/Category	Professional, course
Customer contact	String/Category	mobile, telephone
Customer age	Integer	32,54,72
Customer priority	Integer/Category	1,2 or 3
Type of problem	String/Category	PR1,PR2,PR3
Date of call (month)	Integer	1,5,12
Date of call (week)	String/Category	I, II, III, IV, or V
Date of call (time of call)	String/Category	time1, time2
Time from the last call (days)	Integer	2,4,50,60
Call frequency per week	Integer/Category	1,2,3
Call frequency per day	Integer/Category	1,2,3
Queue waiting time (sec)	Integer	15, 25, 54

Start

Step 1. *Importing a dataset D*

Step 2. *Splitting dataset D into training (80%) – D* and testing (20%) – D^{test} dataset*

Step 3. *Training the model on the training dataset D**

Step 4. *Call class prediction on test case*

Step 5. *Assessment of classification efficiency - comparison with the actual call class*

Step 6. *Storing the classification model*

End

Figure 4. Pseudo code of the prediction algorithm

IV. Results and Discussion

The process used to define hyperparameters in our code involved Grid Search, which is a commonly used technique in machine learning to find the optimal set of hyperparameters for a model. We used Grid Search to find the optimal values for the hyperparameters "penalty", "solver", and "C" of a logistic regression model. The hyperparameters "penalty" had the possible values of "l1" and "l2", "solver" had the possible values of "liblinear", "lbfgs", "sag", "saga", while "C" had a range of values from 0.01 to 1. After defining the grid of hyperparameter values logistic regression model was trained and evaluated for each combination of hyperparameters. The GridSearchCV function from scikit-learn was used, which performs cross-validation for each combination of hyperparameters and returns the set of hyperparameters that result in the best performance. In this way, Grid Search found the best hyperparameters for the model: "C": 0.01, "penalty": "l2", "solver": "lbfgs". These optimal hyperparameters led to improved performance and accuracy for the model.

To evaluate the success of the algorithm, we used the following measures: *Accuracy, Precision, Recall,* and *F1 scores* [28]. The achieved results show that the model has a high accuracy of 95.46%. This means that out of all the predictions made by the model, 95.46% of them were correct. The precision of 92.05% shows that out of all the positive predictions made by the model, 92.05% of them were actually positive. The recall of 99.39% means that the model was able to correctly identify 99.39% of all the actual positive cases. The F1 Score of 0.96 is a metric that combines both precision and recall and gives a single value as a measure of the model's performance.

The Confusion Matrix provides a more detailed view of the model's performance by showing the number of true positive, false positive, true negative, and false negative predictions. In this case, the matrix shows 7132 true positive predictions, 652 false positive predictions, 46 false negative predictions, and 7549 true negative predictions (Table III).

TABLE II. MEASURE OF SUCCESS OF ALGORITHMS FOR CALL CLASSIFICATION

	Accuracy	*Precision*	*Recall*	*F1 Score*
%	95.46	92.05	99.39	0.96

TABLE III. CONFUSION MATRIX

		ACTUAL	
		positive	**negative**
PREDICTED	**positive**	7132	652
	negative	46	7549

V. Conclusion

One of the biggest challenges of managing support call channels is to respond to each incoming customer call uniquely in real time. The job is additionally complicated if we take care that all customers do not react in the same way while waiting in the service queue. In this paper, we have shown that prediction models based on previous customer interactions can predict the customer's propensity to the risk of leaving the queue. To predict the patience of the customer, we ask a key question: "Did the customer wait to be served by the CSR or did he give up?“. This analysis is based on a large data set that includes 76,894 customer calls from a single customer support telephone channel. An important advantage of this research is that we focus on the prediction of call arrival at the individual customer level. This paper demonstrates that by utilizing a specific set of customer attributes as algorithm input, it is feasible to construct a customer patience prediction model with a satisfactory degree of precision.

The classification model has a satisfactory accuracy and opens more questions for further research. First of all, the dynamic attributes of customer behavior created in this way, which are evaluated during each call, are suitable for classifying calls into different classes of call urgency, which could be a topic for future research. If the customer is characterized as impatient for an immediate call, he can have priority when routing the call to a free CSR. Certainly, respecting possible ethical issues, patience would not be the only factor when deciding the priority of answering calls in the practical implementation. But by combining it with other customer profile attributes (especially the priority and the reason for the call) it is possible to avoid "punishing" patient and "rewarding" impatient customer. Second, it would be interesting to apply other classification algorithms and compare their accuracy. The waiting time is the lower limit of the time customer are willing to wait. It would be interesting to use other authors' patience evaluation models, then apply classification algorithms to them, and compare them with our results. Based on the customer's previous interactions, a certain time can be estimated for which the customer must be approached in order not to leave the queue, which will be the subject of future research.

REFERENCES

[1] A. R. Ali, “Intelligent call routing: Optimizing contact center throughput,” *Proc. 11th Int. Work. Multimed. Data Mining, MDMKDD’11 - Held Conjunction with SIGKDD’11*, pp. 34–43, 2011, doi: 10.1145/2237827.2237832.

[2] E. Avdagić-Golub, M. Begović, and A. Kosovac, “Optimization of agent-user matching process using a machine learning algorithms,” *TEM J.*, vol. 9, no. 1, 2020, doi: 10.18421/TEM91-22.

[3] G. Li, J. Z. Huang, and H. Shen, “To Wait or Not to Wait: Two-Way Functional Hazards Model for Understanding Waiting in Call Centers,” *J. Am. Stat. Assoc.*, vol. 113, no. 524, pp. 1503–1514, 2018, doi: 10.1080/01621459.2018.1423985.

[4] R. Ibrahim, H. Ye, P. L’Ecuyer, and H. Shen, “Modeling and forecasting call center arrivals: A literature survey and a case study,” *Int. J. Forecast.*, vol. 32, no. 3, pp. 865–874, 2016, doi: 10.1016/j.ijforecast.2015.11.012.

[5] A. D. Kelleher, J Namee B, *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. The MIT Press, 2015. [Online]. Available: https://www.ptonline.com/articles/how-to-get-better-mfi-results

[6] M. Saberi, O. Khadeer Hussain, and E. Chang, “Past, present and future of contact centers: a literature review,” *Bus. Process Manag. J.*, vol. 23, no. 3, pp. 574–597, 2017, doi: 10.1108/BPMJ-02-2015-0018.

[7] V. M. Zeynep Aksin, Mor Armony, “The Modern Call Center: A MultiDisciplinary Perspective on Operations Management Research,” *Prod. Oper. Manag. Soc.*, vol. 16, no. 6, pp. 657–664, 2007, doi: 10.3401/poms.

[8] B. Legros, O. Jouini, and Y. Dallery, “A flexible architecture for call centers with skill-based routing,” *Int. J. Prod. Econ.*, vol. 159, pp. 192–207, 2015, doi: 10.1016/j.ijpe.2014.09.025.

[9] A. Senderovich, M. Weidlich, A. Gal, and A. Mandelbaum, “Queue mining for delay prediction in multi-class service processes,” *Inf. Syst.*, vol. 53, pp. 278–295, 2015, doi: 10.1016/j.is.2015.03.010.

[10] F. Iravani and B. Balcioǧlu, “On priority queues with impatient customers,” *Queueing Syst.*, vol. 58, no. 4, pp. 239–260, 2008, doi: 10.1007/s11134-008-9069-6.

[11] A. Dudin, C. Kim, O. Dudina, and S. Dudin, “Multi-server queueing system with a generalized phase-type service time distribution as a model of call center with a call-back option,” *Ann. Oper. Res.*, vol. 239, no. 2, pp. 401–428, 2016, doi: 10.1007/s10479-014-1626-2.

[12] O. Dudina and S. Dudin, “Queueing system MAP/M/N/N + K operating in random environment as a model of call center,” *Commun. Comput. Inf. Sci.*, vol. 356 CCIS, pp. 83–92, 2013, doi: 10.1007/978-3-642-35980-4_10.

[13] M. G. Nosova and M. V. Degtyareva, “Building and Analysis of the Contact Center Model As a Queuing System With Multi-Level Ivr and Impatient Applications,” *Вестник Алтайской Академии Экономики И Права*, no. №9 2019, pp. 101–106, 2019, doi: 10.17513/vaael.705.

[14] M. Abhishek Pandharipande and S. K. Kopparapu, “A novel approach to identify problematic call center conversations,” *JCSSE 2012 - 9th Int. Jt. Conf. Comput. Sci. Softw. Eng.*, no. April, pp. 1–5, 2012, doi: 10.1109/JCSSE.2012.6261805.

[15] J. Chatterjee, A. Saxena, and G. Vyas, “An automatic and robust system for identification of problematic call centre conversations,” *Int. Conf. Micro-Electronics Telecommun. Eng. ICMETE*, pp. 325–330, 2016, doi: 10.1109/ICMETE.2016.48.

[16] V. U. of W. Rafiq A. Mohammed1, “Using Personalized Model to Predict Traffic Jam in Inbound Call Center,” *EAI Endorsed Trans. Scalable Inf. Syst.*, 2017, doi: http://dx.doi.org/10.4108/eai.18-1-2017.152101.

[17] R. A. Mohammed and P. Pang, “Agent personalized call center traffic prediction and call distribution,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7063 LNCS, no. PART 2, pp. 1–10, 2011, doi: 10.1007/978-3-642-24958-7_1.

[18] S. Moazeni and R. Andrade, “A Data-Driven Approach to Predict an Individual Customer’s Call Arrival in Multichannel Customer Support Centers,” *Proc. - 2018 IEEE Int. Congr. Big Data, BigData Congr. 2018 - Part 2018 IEEE World Congr. Serv.*, pp. 66–73, 2018, doi: 10.1109/BigDataCongress.2018.00016.

[19] P. D. Feigin, “Analysis of Customer Patience in a Bank Call Center,” 2006.

[20] G. Koole and A. Pot, “An overview of routing and staffing algorithms in multi-skill customer contact centers,” *Submitt. Publ.*, pp. 1–42, 2005, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.1714&rep=rep1&type=pdf

[21] J. Chen, J. Dong, and P. Shi, “A survey on skill-based routing with applications to service operations management,” *Queueing Syst.*, vol. 96, no. 1–2, pp. 53–82, 2020, doi: 10.1007/s11134-020-09669-5.

[22] X. Sun, “Multi-attribute intelligent queueing method for onboard call centers,” *Cluster Comput.*, vol. 22, pp. 5207–

5215, 2019, doi: 10.1007/s10586-017-1173-0.

[23] E. Avdagic-Golub, M. Begovic, S. Causevic, and A. Kosovac, "Profiling Contact Center Customers for Optimization of Call Routing Using Data Mining Techniques," *2021 20th Int. Symp. INFOTEH-JAHORINA, INFOTEH 2021 - Proc.*, no. March, pp. 17–19, 2021, doi: 10.1109/INFOTEH51037.2021.9400671.

[24] F. Marisa, S. S. S. Ahmad, Z. I. M. Yusof, Fachrudin, and T. M. A. Aziz, "Segmentation model of customer lifetime value in Small and Medium Enterprise (SMEs) using K-Means Clustering and LRFM model," *Int. J. Integr. Eng.*, vol. 11, no. 3, pp. 169–180, 2019, doi: 10.30880/ijie.2019.11.03.018.

[25] L. Brown *et al.*, "Statistical analysis of a telephone call center: A queueing-science perspective," *J. Am. Stat. Assoc.*, vol. 100, no. 469, pp. 36–50, 2005, doi: 10.1198/016214504000001808.

[26] M. Thiongane, W. Chan, P. Ecuyer, M. Thiongane, W. Chan, and E. Waiting, "Waiting time predictors for multi-skill call centers," in *2015 Winter Simulation Conference, Huntington Beach, United States*, 2015.

[27] A. I. Kyritsis and M. Deriaz, "A machine learning approach to waiting time prediction in queueing scenarios," *Proc. - 2019 2nd Int. Conf. Artif. Intell. Ind. AI4I 2019*, pp. 17–21, 2019, doi: 10.1109/AI4I46381.2019.00013.

[28] C. I. Eke, A. A. Norman, L. Shuib, and H. F. Nweke, "A Survey of User Profiling: State-of-the-Art, Challenges, and Solutions," *IEEE Access*, vol. 7, pp. 144907–144924, 2019, doi: 10.1109/ACCESS.2019.2944243.