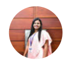
A Machine Learning Approach to Predict the Personalized Next Payment Date of An Online Payment Platform

Conference Paper · December 2022

DOI: 10.1109/ICAC57685.2022.10025194

Chamodi Karunathunge
Sri Lanka Institute of Information Technology



M.G.N.M. Pemadasa
Sri Lanka Institute of Information Technology

A Machine Learning Approach To Predict The Personalized Next Payment Date Of An Online Payment Platform.

L. C. R. Karunathunge
Department of Information Technology
Sri Lanka Institute of Information Technology, Sri Lanka
it19118864@my.sliit.lk

B. N. Dewapura
Department of Information Technology
Sri Lanka Institute of Information Technology, Sri Lanka
it19176734@my.sliit.lk

V. A. S. Perera
Department of Information Technology
Sri Lanka Institute of Information Technology, Sri Lanka
it19145662@my.sliit.lk

G. P. R. A. Kavirathne
Department of Information Technology
Sri Lanka Institute of Information Technology, Sri Lanka
it19113364@my.sliit.lk

A. Karunasena
Department of Information Technology
Sri Lanka Institute of Information Technology, Sri Lanka
anuradha.k@sliifromlk

M. G. N. Pemadasa
Departtype-wisenformation Technology
Sri Lanka Institute of Information Technology, Sri Lanka
nadeesa.p@sliit.lk

***Abstract*—Use of digital payments has risen exponentially in the recent past especially due to the COVID-19 pandemic. This is because online payment methods offer many benefits in performing their day-to-day transactions and paying utility bills such as electricity bills, water bills, telephone bills and etc. Knowing when a consumer will perform a specific online transaction, or bill payment is beneficial to an online payment platform to plan marketing campaigns since targeted marketing has become very prevalent nowadays. However, predicting this is not an easy task since thousands of transactions are happening in each and every minute of an online payment platform. This paper presents the results of a study that investigated predicting the customer personalized, utility bill payment type wise next payment date of a financial company in Sri Lanka by using machine learning techniques. This is accomplished by analyzing not only online transaction history but also customer characteristics and a holiday calendar which is specific to Sri Lanka. At the end of the study, it was identified that XGBoost Regressor is the most suitable machine learning algorithm, etc deal with this scenario which provided 91.02% accuracy. These predictions will be used for sending personalized reminders and discount offers to customers without sending general common notifications when they are planning to do an online payment. Such reminders and offers will be notified on the mobile devices of the customers and, ultimately both customers and the business owners will be benefited by this.**

***Index Terms*—online payment platforms, next payment date, predict, machine learning**

I. Introduction

The global payments industry is rapidly evolving and is constantly in flux due to the introduction of new payment methods, mergers and acquisitions, and new technology [1]. In recent years, digital payment solutions have expedited the evolution of the payment ecosystem. Due to the growth and expansion of the global online payments network [2] and technology, businesses that develop payment platforms play an essential role in the payment industry. Many are partnering with established financial institutions to satisfy customer and merchant preferences with innovative ways to increase digital money transfer in society by re-imagining the customer experience and facilitating business owners' operations. Nowadays payment processing is more than just facilitating the transfer of funds [1]. During the COVID-19 pandemic, online payment platforms expanded, resulting in a rise in digital payments. [3] According to a global survey conducted by The World Bank Group across 109 countries in the world, more than 80% of respondents stated that the COVID-19 pandemic has increased the need for fin-tech and digital transformation [4].

Online payment platforms have become popular since they offer numerous benefits. According to a study by S. Yakean [5], a cashless society can reduce corruption because online financial transactions are more transparent, and crimes can be reduced as there's no cash to steal. Furthermore, daily transactions and utility bill payments such as electricity bills, water bills, etc. can be made from anywhere and at any time, saving time and energy over going to the bank. The use of online payment platforms has aided in the reduction of the global COVID 19 epidemic as well as the fuel crisis.

Despite the advantages, online payment platforms have disadvantages associated with them. Due to the fact that online payment platform apps require a smartphone or tablet to process COVID-19, clients cannot transact with merchants who lack these facilities. According to S. Yakean [5], this tendency may encourage more cybercriminals to target e-payments. In addition, online payments reduce privacy as their payment

979-8-3503-9809-0/22/$31.00

information may surface in unexpected ways [5]. Despite some restrictions, the benefits of online payment systems and the digitization of payments outweigh their drawbacks.

Financial institutions can gain many benefits by redefining the customer experience, as payment processing is now more than just facilitating the transfer of funds. One such possible way of redefining the customer experience is knowing when a consumer will perform a specific online transaction, such as a bill payment and reminding customers. Reminding customers to pay bills by predicting their next payment date for a specific transaction type is a very convenient way for customers to remember to pay their bills through a particular online payment platform. From the perspective of the financial institution, this would be helpful for target marketing which has become very popular in recent years. The objective of this research is to provide a personalized solution to predict the utility bill payment type-wise next payment date for a particular customer by using machine learning.

Machine learning techniques can predict the customer's personalized utility bill payment type or transaction type-wise next payment date for the benefit of both the customer and the business. This can be approached by analyzing not only the online transaction history of customers of an online payment platform but also their customer characteristics and the holidays associated with transaction days. Without providing generalized common notifications, these predictions can be utilized to send personalized reminders and discount offers to clients who are intending to make an online payment. Such reminders and offers will be notified on the mobile devices of the customers and, ultimately both customers and the business owners will be benefited by this. [1]

II. Related Works

This chapter explains the background study of the research and the works that are most closely related to the challenges addressed in this paper. It was identified that there was few research done to predict the next payment date of a customer. For example, as per M. Droomer and J. Bekkers's research on "Using machine learning to predict the next purchase date for an individual retail customer" [6], aimed at developing a predictor that predicts when an individual would purchase a specific product again. This has been done by analyzing the past purchasing transaction behaviours of individual customers. The basic idea behind this research study [6] is that a customer purchases items from a retail store which records the customer's purchase history. To predict the next payment date (NPD) machine learning techniques are applied to sales data to predict when a customer will buy a specific product again. This information is used by the store's marketing team to target a specific individual at the right time. To predict the NPD the researcher has tried several machine learning techniques on the "Instacart online grocery shopping" dataset of 2017 has been used. Linear Regression has only looked at the sequence of one user-product pair, whereas the other two techniques, Artificial Neural Networks (ANN) and Extreme Gradient Boosting, have looked at the big picture by using all of the data from all of the user-product pairs [6]. In the end, the researchers evaluated the models and chosen the Neural Network model as the best model with the highest accuracy as the NPD Predictor.

Another research study [7] has been conducted to predict the payment date of an invoice when it gets created in the system. In addition, the research divides the invoice into different buckets based on the expected payment date. The invoices dataset used in the research contained past payment information and buying patterns. The machine learning models were built using Random Forest Regressor, Decision Tree Regressor, XGB Regressor and Linear Regression algorithms based on previous payment patterns to predict when a customer will make a payment for an invoice. Based on the predicted payment date, the model also predicted which ageing bucket the invoice would fall into. The different buckets were as below.

- 0-15 days
- 16-30 days
- 31-45 days
- 46-60 days
- Greater than 60 days

In the research study [8], the authors have used six months of behavioural data to predict customers' first purchase date in the next three months. If there are no purchase history records for a particular customer, they have predicted the next purchase date of them as well. In this research, the authors have built classification models where the predicted class label would be one from the following :

- 0–20: Customers that will purchase in 0–20 days
- 21–49: Customers that will purchase in 21–49 days
- Greater than or equal 50: Customers that will purchase in more than 50 days

After building several machine learning models, Naive Bayes has selected as the best-performing model which had 64% accuracy.

Another research study [9] which has been conducted by the Department of Computer Science of Vietnam National University, has focused on predicting the next purchase item in "Vo Lam Truyen Ky Mobile" (JXM) which is a Vietnam's four major game publishers' (VNG) hot game. In this research study [9], the researchers have divided the data set into 25 clusters by using K-Means Clustering and with the top 5 most popular items taking account for nearly 80% of all transactions in each cluster according to date. Furthermore, using the ARIMAX model, they predicted 5 next purchase items, accounting for roughly 70% of the total number of purchased items to date. The proposed method outperformed Collaborative Filtering in terms of results with an accuracy of 22%. The findings of the study were used to recommend products that are relevant to users, improve user experience, and potentially increase revenue.

According to a background study and literature review conducted on previous research studies relevant to the next payment date prediction, some research gaps were identified. As previously stated, although there were few research done to predict customers' next payment date and there is a limitation of research done to predict both customers next payment date and utility bill type. Furthermore, research studies such as [6], [7], [8] and [9], are not specifically related to an online payment platform. Rather, most of them have done either in relation to the supermarket, online shopping malls or just customer day-to-day transaction date predictions for banks. Hence, this research has been undertaken to develop a complete solution for an online utility bill payment platform that predict both the customer's next payment date and the type of utility bill by using a machine learning approach.

III. Methodology

In order to meet the objective of this research, there were a few steps to be followed as displayed on Fig. 1. This section discusses those steps that need to be followed to carry out in detail. This section further discusses the environments, techniques, requirements, tools, and technologies that are used to achieve each milestone.

A. Data Collection

This research study was done using data gathered from one of the leading digital payment platforms in Sri Lanka. Basic customer characteristics (customer age, gender, district, employment status, registered year and etc.) and past transactional data of customers who have done at least 10 utility bill payment transactions during 2020 and 2021 have are captured in the dataset. A total of 215,208 data records were collected under above above-mentioned conditions. In addition to that, a calendar data set was implemented by considering all the weekends, bank, mercantile and public holidays.

B. Data Preprocessing & Model Building

Initially, data related to customer transactions, customer characteristics, and the implemented holiday calendar were uploaded separately to a Python notebook.

Customer ID, Utility Bill Payment Type, Transaction ID, and Transaction were included in the past transactions data set, as indicated in the table below (Table I).

TABLE I: Past Transaction Data (Before Preprocessing)

Customer ID	Utility Bill Type	Transaction ID	Transaction Date
001	CEB	222158	2021-02-18
002	Telecom	334586	2021-03-02
003	Dialog	342688	2021-03-21
001	Water Bill	344924	2021-03-26
003	CEB	412654	202employment
004	Te, etc.25876	2021-04-24	
005	Water Bill	485066	2021-05-09

The customer characteristics data set included basic information about each consumer. Table II illustrates the sample format of these data for only few of the variable features.

Holiday Calendar was implemented by taking Sri Lankan holidays into account, as this data set was also sourced from a digital payment platform in Sri Lanka. The following Table III is a sample data set for the implemented Holiday Calendar.

As the initial step of data pre-processing, categorical data were handled by transforming them to numerical values using

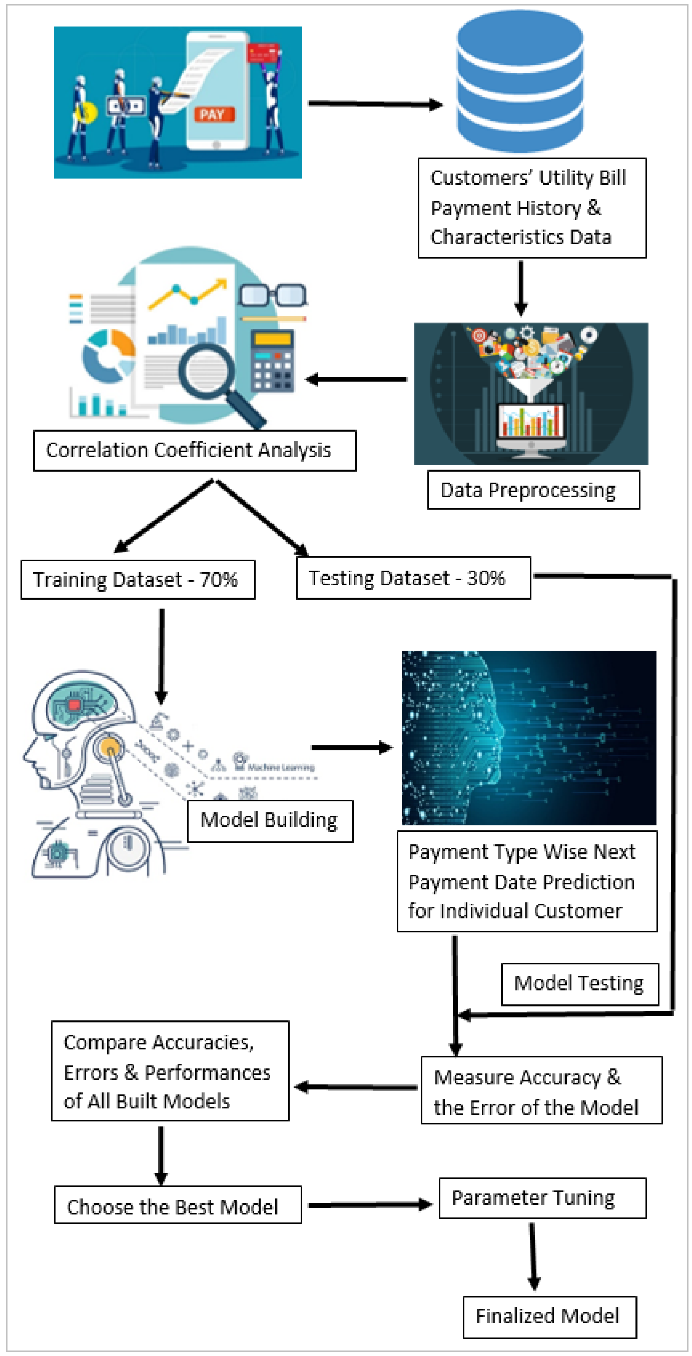


Fig. 1: Methodology Diagram

TABLE II: Customer Characteristics Data (Before Preprocessing)

Cus_ID	Gender	Age	District	Reg_Year	Employment_Status
001	Male	32	Galle	2017	Employed
002	Male	28	Colombo	2017	Unemployed
003	Female	56	Kandy	2018	Employed
004	Male	19	Colombo	2018	Unemployed
005	Female	64	Colombo	2020	Unemployed

ordinal encoding methods and guided encoding methods. Ordinal encoding converts each categorical value into an integer

TABLE III: Holiday Calendar (Before Preprocessing)

Date	Day_Name	Bank	Mercantile	Public	Weekend
2021-04-08	Thursday	0	0	0	0
2021-04-09	Friday	0	0	0	0
2021-04-10	Saturday	0	0	0	1
2021-04-11	Sunday	0	0	0	1
2021-04-12	Monday	0	0	0	0
2021-04-13	Tuesday	1	1	1	0
2021-04-14	Wednesday	1	1	1	0
2021-04-15	Thursday	0	0	0	0
2021-04-16	Friday	0	0	0	0

value, whereas guided encoding replaces the categorical value with a specified value. The "Null" and garbage values were, then, handled by assigning unique values and removing rows if the number of rows was negligible.

Consequently, it was utilized to transform the transaction-wise data set (Table I) into a format that could be used to predict the utility payment type-wise next payment date of customers. In order to do that, the data set was grouped by "Customer ID" and "utility bill payment type".

Next, it was attempted to derive the target variable and other insightful supporting variables that aid in determining the next payment date. As the first step to do so, the "Number of Days Between Consecutive Transactions" was derived from the transaction data set. This was calculated by considering the utility payment type of each customer. After that, in order to have a limit to the customer and utility bill payment type-wise transactions, only the last 7 transactions were chosen to proceed with the model building. This decision was made because some customers have a huge number of transactions for a particular utility bill type, while others have a very little number. Consequently, using only the most recent seven transactions helps to resolve this data imbalance issue. The holiday Calendar was combined next to the above transaction data set by considering the "Transaction Date" of each transaction. Following the above, the availability of multiple rows for each customer needed to be handled. In order to do that, one row for each utility bill payment of customer was derived as below. Furthermore "Transaction Date" and "Transaction ID" were removed from the data frame since it is no longer necessary.

Another two variables were derived by considering the "Number of Days Between Consecutive Transactions" as below.

- Max Days – This will be the highest value of "Number of Days Between Consecutive Transactions" for each utility type of each customer.
- Min Days – This will be the lowest value of "Number of Days Between Consecutive Transactions" for each utility type of each customer.
- Average Days - This will be the average value of "Number of Days Between Consecutive Transactions" for each utility type of each customer.

After the above process, the transaction data set has only one row for each utility bill type of customer combined with the Holiday Calendar and other derived variables. Table IV shows a few of the features that resulted after the process explained above.

- ID – Customer ID
- UT – Utility Bill Type
- t1-7–Number of Days Between Consecutive Transactions
- h1-7 – Holiday Type (1/0)
- Mx – Max Days Mn - Min Days
- A – Average Days

TABLE IV: Transaction Data Set with derived columns and holidays (After Preprocessing)

ID	UT	t1	t2	t3	t4-7	h1	h2-7	Mx	Mn	A
1	21	N/A	6	8	...	0	...	10	3	6.7
1	36	N/A	28	24	...	0	...	52	10	32.7
1	14	N/A	15	21	...	0	...	21	9	17.2
2	36	N/A	32	36	...	0	...	42	11	33.8
2	6	N/A	23	28	...	1	...	30	6	26.8
3	21	N/A	4	12	...	0	...	14	4	8.7

As can be seen in the Table IV, the first transaction of each row does not have a value for "Number of Days Between Consecutive Transactions". Therefore, that column should be removed from the data frame. As the next step, customer characteristics data are combined with the transaction data set by considering the "Customer ID" of each table. "Number of Days Between Consecutive Transactions" value relevant to the last transaction (t7 - 7th transaction) which means **number of days left for the next payment date** will be the **target variable** of this data set.

After the above process of data cleaning and data transformation, data analysis was done to explore data and a correlation coefficient analysis was done to measure the strength of the relation between variables and determine the variables which have an impact on the prediction. A correlation coefficient greater than zero indicates a positive relationship, whereas a correlation coefficient value less than zero indicates a negative relationship. A value of zero indicates that the two variables being compared have no correlation [10]. To determine the most appropriate variables, the absolute value of the correlations with the target variable was calculated, and all variables with a value greater than 0.70 were eliminated.

After creating the features and datasets, the models were developed and evaluated. The initial strategy consisted of predicting the utility bill type-wise next payment date using classification techniques. As previously stated, the target variable was *"Number of days left for the next payment date "* (t7 - number of days between consecutive transactions value relevant to the 7th transaction). Random Forest Classifier, Decision Tree Classifier, KNeighbors Classifier and Bagging Classifier were used when building the model. However, they all ended up getting low accuracies as depicted on Table V.

As a result, it was decided to use regression algorithms as the next step. Before fitting to the model, all the days count in finalized data set as per Table IV needed to be converted to weeks as the target variable must have continuous values when fitting to the model. So, as per in Table IV, "t2", "t3", "t4"," t5", "t6", "t7", "Max Days", "Min Days" and "Average

Days" values were transformed to weeks by dividing all the day counts from 7.

The models were built using the XGBoost Regressor, Random Forest Regressor, and Decision Tree Regressor algorithms. The target variable was *"Number of weeks left for the next payment date"* (t7 days count has transformed to weeks here). Initially, the finalized data set was split to two sets as training data set (70%) and testing data set (30%). Then the training data set was used to fit to the algorithms and the accuracies and errors were calculated (Table VI). Then, the algorithm which has the highest accuracy, highest performances and the lowest error was chosen as the best algorithm. According to the above criterion, XGBoost Regressor was selected as the best algorithm with 87.4% accuracy to predict the utility bill type wise next payment date (No of weeks left for the next payment date). Also, its execution speed and model performance were an added advantage.

After choosing the XGBoost Regressor as the best algorithm, variables and parameters were tuned to increase the accuracy, and performance and to decrease the error before finalizing the model. Grid Search was used to tune the hyper-parameters of the XGBoost algorithm here. After hyper-parameter tuning, the accuracy was 91.02% and the mean absolute error was 0.71.

The XGBoost model was implemented by using the XGBoost package in Python with the following parameters to achieve the highest accuracy and lowest error.

- n_estimators (Number of trees in the ensemble) = 1400
- max_depth (Maximum depth of each tree) = 5
- eta (Learning rate) = 0.1
- subsample (Number of samples rows used) = 1.0
- colsample_bytree (Number of features) = 1.0

After predicting the number of weeks left for the next utility bill payment, the exact next payment date for a specific utility bill payment type can be determined for each customer by adding the predicted "Weeks left for the next payment date" value to the last transaction date of a particular utility bill payment type as below.

$$NPD = LPD + (PWC * 7)$$

NPD - Next Paymnet Date
LPD - Last Payment Date
PWC - Predicted Weeks Count

IV. Results & Discussions

This section will be discussing the results obtained from the machine learning models implemented using the methodology discussed above.

As stated previously, several classification techniques were initially implemented, and the accuracies obtained for each algorithm are listed in Table V below. Evidently, the accuracy of the algorithms used to predict the next payment date was insufficient to proceed further.

Considering the implemented machine learning techniques for regression, the following accuracies and errors were obtained, as shown in Table VI. After measuring the accuracy, error, and performance of each method, the model with the highest accuracy, highest performance, and the lowest error rate was selected as the optimal algorithm. Accordingly, XGBoost Regressor was identified as the best algorithm to predict the utility bill type-wise next payment date. The accuracy was 91.02% and the mean absolute error was 0.71 for the finalized XGBoost Regression algorithm which was selected as the most suitable algorithm.

TABLE V: Results of Classification Algorithms

Classification Algorithm	Accuracy	Precision	Recall	F1 Score
Decision Tree Classifier	51.35%	43.41%	64.20%	51.80%
Random Forest Classifier	45.19%	48.62%	51.47%	50.00%
K Neighbors Classifier	37.86%	52.94%	36.84%	43.45%
Bagging Classifier	24.25%	36.05%	44.29%	39.75%

TABLE VI: Results of Regression Algorithms

Regression Algorithm	Accuracy	R2	Error
XGBoost Regressor	91.02%	0.88	0.71
Random Forest Regressor	73.12%	0.73	1.5
Decision Tree Regressor	41.48%	0.46	1.68

XGBoost uses decision trees as its primary learners, merging numerous weak learners to produce a powerful learner. As a result, it is known as an ensemble learning method since the final prediction is based on the output of multiple models. XGBoost is the optimal combination of software and hardware optimization strategies for producing better results with fewer computer resources and in less time. As mentioned earlier, XGBoost has numerous benefits. XGBoost is very flexible and supports parallel processing. In addition, regularization is allowed, and cross-validation can be performed after each iteration [**?**].

Furthermore, the accuracy and performance of the models implemented in this research study was improved than the previously stated research [6], [7], [8] and [9] while keeping errors to a minimum. In addition, in other research materials, only the customer transaction history was used to build the models. But this research has been done using a holiday calendar associated to the transaction days as well as customer characteristics that have an impact on predicting the next utility bill payment date and type.

As mentioned earlier, different models were implemented using different machine learning algorithms to find the most suitable model algorithm. Furthermore, real data obtained from a financial company in Sri Lanka was used to build the models. After analyzing all the models implemented, the XG Boost Regression algorithm was chosen over other algorithms as the best model based on accuracy, performance, and error rates, whereas in previous research [6], [7], [8] and [9] the best models were Linear Regression, ANN, Nave Bayes, and ARIMAX algorithms.

Moreover, developing a new machine learning model to predict personalized utility bill payment type wise next payment date by bridging the above-mentioned research gaps of the previous research studies [6], [7], [8] and [9] benefits not only businesses for marketing and make customers stick to the application, but also customers as well for managing their own bill payments. As a result, establishing accurate personalized utility bill payment type-wise next payment date prediction model will become a pillar of an effective marketing plan to persuade customers to do their utility bill payments through these online payment platforms which will highly benefit the online payment platform owners, business as well as the users. After identifying the user and utility bill type-wise next payment date, the online payment platforms' management can make a necessary marketing plan to notify customers to persuade for bill payments as well as the merchants for creating offers and discount periods. This is the first research attempt to implement a personalized utility bill payment type-wise next payment date prediction model for an online utility payment platform.

V. Limitations & Future Works

One of the limitations of this research is that data for the years 2020 and 2021 were acquired from an online payment platform. The impact of the COVID pandemic and the economic crisis in Sri Lanka is reflected in those data. As a result, the behaviours of customers and the patterns of their transactions can be changed over time. Therefore, it is best to train the model using data from more recent times.

Furthermore, for future work, a model can be built using a sequence-based method and a Neural Network as the above prediction model was developed solely based on the non-sequenced method.

VI. Conclusion

It was possible to predict the customer personalized, utility bill payment type wise next payment date. This paper has discussed how to predict that next payment date by using past customer transaction history, customer characteristics as well as holiday data which has an impact to the next payment date prediction to a particular customer.

It was discovered that the XGBoost Regression is the most suitable approach for the prediction out of many regression and classification algorithms. The accuracy of XGBoost Regression model was 91.02% and the mean absolute error was 0.71. This prediction can be used to send personalized reminders and discount offers to customers who are planning to make an online payment without sending generic notifications. These reminders and promotions can be sent to customers' mobile devices, which will ultimately benefit both customers and business owners.

References

[1] Varone, "The Payments Industry Landscape: What Does It Look Like Today?," Cardknox, May 06, 2020. https://www.cardknox.com/white-papers/payments-industry-landscape/ (accessed Sep. 05, 2022).

[2] "Digital payments shaping future of E-Commerce in Sri Lanka - News — PayMedia," www.paymedia.lk. http://www.paymedia.lk/news/5-digital-payments-shaping-future-of-e-commerce-in-sri-lanka/ (accessed Sep. 05, 2022).

[3] S. Jesuthasan and N. Umakanth, "Impact of Behavioural Intention on E-Wallet Usage During Covid-19 Period: A Study from Sri Lanka," *Sri Lanka Journal of Marketing*, vol. 7, no. 2, p. 24, Aug. 2021, doi: 10.4038/sljmuok.v7i2.63.

[4] F. Paul,Saal,Matthew,Sarkar,Arpita Erik H. B. ,Natarajan,Harish,Heffernan,Robert, "World Bank Group Global Market Survey : Digital Technology and the Future of Finance," World Bank. http://documents.worldbank.org/curated/en/099735404212273637 (accessed Sep. 05, 2022).

[5] S. YAKEAN, "Advantages and Disadvantages of a Cashless System in Thailand during the COVID-19 Pandemic," *The Journal of Asian Finance, Economics and Business*, vol. 7, no. 12, pp. 385–388, Dec. 2020, doi: 10.13106/jafeb.2020.vol7.no12.385.

[6] M. Droomer and J. Bekker, "Using machine learning to predict the next purchase date for an individual retail customer," *S. Afr. J. Ind. Eng.*, vol. 31, no. 3, pp. 69–82, 2020.

[7] "Machine Learning Model to predict the payment date of an invoice when it gets created in the system.— PythonRepo," pythonrepo.com. https://pythonrepo.com/repo/SkywalkerHub-Payment-Date-Prediction-python-machine-learning (accessed Sep. 05, 2022).

[8] B. Karaman, "Predicting Next Purchase Day," Medium, Sep. 15, 2019. https://towardsdatascience.com/predicting-next-purchase-day-15fae5548027 (accessed Sep. 05, 2022).

[9] T. Nguyen, T. Le, and B. Le, "Predicting next purchase item on JXM game by K-means clustering and ARIMAX model," *2020 7th NAFOSTED Conference on Information and Computer Science (NICS)*, 2020.

[10] S. Nickolas, "What does it mean if the correlation coefficient is positive, negative, or zero?," Investopedia, May 31, 2021. https://www.investopedia.com/ask/answers/032515/what-does-it-mean-if-correlation-coefficient-positive-negative-or-zero.asp (accessed Sep. 05, 2022).

[11] "XGBoost an efficient implementation of gradient boosting," datascience.foundation. https://datascience.foundation/datatalk/xgboost-an-efficient-implementation-of-gradient-boosting (accessed Sep. 05, 2022).