Proceedings of Sixth International Congress on Information and Communication Technology.
Lecture Notes in Networks and Systems, Volume 217. Springer, Singapore
DOI:10.1007/978-981-16-2102-4_15

Using client's characteristics and their financial products to predict their usage of banking electronic channels

Hessa Alsanousi[1,2][0000-0001-5820-3781], Najla Albarrak[1,2][0000-0003-1502-7268], Irene Moulitsas[2][0000-0003-0947-9495] and Salvatore Filippone[2,3][0000-0002-2860-8566]

[1] Central Bank of Kuwait, P. O. Box: 526, Safat 13006, Kuwait
[2] Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK
[3] University of Rome Tor Vergata, Via Cracovia, 50, 00133 Roma RM, Italy
h.alsanousi@cranfield.ac.uk
n.albarrak@cranfield.ac.uk
i.moulitsas@cranfield.ac.uk
salvatore.filippone@cranfield.ac.uk

Abstract. Technology innovation and its impact on the progress of electronic banking establishes the requirement for this research regarding customer demographics, their financial portfolio and banking preferences. In this research, banking financial data was collected from three Kuwaiti banks. The data included usage information in all electronic banking channels for each customer, their characteristics and financial portfolio. The aim of this study is to predict the customer use of electronic channels, treated as dependent variables, considering individual customer information, that is treated as independent variables. To bring the most benefit to bankers and financial analysts, machine learning techniques (ML), specifically multinomial logistic regression, were used to deal with the data from cleaning to analysing. The results disclosed that banks can determine the preferred electronic banking channel for each of their customers by knowing some information about their characteristics and financial product portfolio.

Keywords: Electronic Banking, Mobile Banking, Call Centre, Online Banking, Financial Services, FinTech (Financial Technology), Machine learning, Multinomial Logistic Regression, Banking Industry, MATLAB, Kuwait Market, Islamic Banks, Conventional Banks

1 Introduction

One of the essential sources of operation in the banking sector is the electronic banking. Through the efficient use of electronic channels (call centre platforms, online websites, mobile applications) banks will increase their revenue and decrease the cost of serving their customers. This will enhance bank's efficiency and profitability. Knowing the appropriate online channel for each customer, based on their characteristics and their

Published by Springer. This is the Author Accepted Manuscript issued with: Creative Commons Attribution Non-Commercial License (CC:BY:NC 4.0).
The final published version (version of record) is available online at DOI:10.1007/978-981-16-2102-4_15. Please refer to any applicable publisher terms of use.

banking portfolio, will help increase customer satisfaction. Eventually, resulting in a continuous and productive relationship where customers will demand more banking services and products. Not only will this increase bank's revenue, but it will also decrease costs (e.g., advertisements cost) by letting the banks determine the appropriate electronic channel for their customers. This will assist banks to advertise for the use of their communication channels to the right customers: the customers interested in each channel. The main motivation of this study is to assistance banks on how to deal with the rapid growth of technology, mobile applications, in particular. By understanding their clients' preferences and to what limit they should go with these new technologies. This will improve bank's management strategies for better allocation of their expenses and capital expenditure and eventually robust their profit. Today's financial world witnessed the evolution of banking data which allowed for the collection of substantial amounts of relevant data for this study.

2 Literature Review

Developments in technology have led to more banks offering new banking services while effectively cutting costs and enhancing customer satisfaction [1].One of the most recent paper indicated that in order to increase mobile banking adoption intention, banks could focus on accessibility, operation, benefit and post-benefit convenience[2].

Limitations and restrictions has been discussed in term of accessibility of the use of internet banking which caused a shift toward the innovation of mobile banking [3]. They did a study on 732 banking customers given their gender, age, and educational level. The results demonstrated that the client age has insignificant impact on the use of any more advanced technology communication channel, so they are willing to accept and adopt any new technology. The outcomes also showed that the bank's clients are fortunate to pay for improved services while benefiting from its advantages. The launch of new technology does not affect their trust in using such technology.

A study was conducted to discover and recognize the factors that could affect the usage of mobile banking application in Kingdom of Saudi Arabia [5]. They used 10 variables related to customers satisfaction and their usage behaviour. Findings showed that performance expectancy, facilitating condition, price value, hedonic motivation, habit, system quality and service quality have significant impact on usage behaviour. Outcomes also showed that learning effort expectancy, social influence and information quality have a little effect on the usage.

The differences of reactions and responses to the use of mobile applications has been considered for banking services in two countries with different stage of development which are Brazil and United States [4]. They used some variables as determinant factors of mobile banking use. The results indicate that trust and ease of use are significant elements to understand the use of mobile banking application in both countries. Results also showed that social effect has a relevant influence on mobile usage in brazil while have no influence in united states.

SPSS and AMOS has been utilized in Jordan as methods to explain the slow uptake of e-banking[6]. In their research, they collected 328 completed surveys distributed to

individuals in Amman that revealed several major factors causing the delay in adopting e-banking. These factors were ease of use, perceived usefulness, security, and reasonable price.

One of the studies focused on investigating why older people are less likely to adopt mobile banking in the United Kingdom through a comparative analysis on Google Scholar and Scopus [7]. The work suggested that future studies is required to provide insight into the factors leading to the adoption of smartphones for mobile banking. This and the above-mentioned studies guided the research used in our study.

3 Methodology

The evolution in technology combined with the rapid growth of banking data led to this study's usage of machine learning techniques to analyse the data. Bankers and financial analysts will not fully benefit from the large volume of available data without the use of machine learning techniques to clean, process, and generate predictive analysis. A supervised machine learning algorithm was used to solve a classification problem by applying a multinomial logistic regression model.

When there are more than two possible outcomes, nominal responses, multinomial logistic regression can be used to calculate the probabilities of various possible outcomes, dependent variables. Multinomial logistic regression (MLR) is a classification model based on binary logistic regression model to solve a multiclass problem. It predicts the probabilities of multiple possible outcomes of a categorical variable by using a set of independent variables[8]. In this research, the dependent variables are the electronic channels.

We focused in this study on three electronic channels (bank's website (internet), call centre and mobile) as a dependent variable; in multinomial logistic regression you must choose one dependent variable as a reference category. In this study, "mobile" is the reference category. The probability of using any of the other online channels is compared to the probability of using "mobile". Rather than running individual regressions and then interpreting and comparing the outcomes. The MATLAB software environment and language were used [9]. This type of regression evaluated parameters simultaneously, resulting in efficiency. This method outputs several logistic regression models that make specific comparisons of the response categories. Assuming there are j-categories of the response variable, then, the model comprises of j-1 logit equations which fit at the same time [10].

MLR classifies d-dimensional real-valued input vectors x ϵ Rd into one of the k outcomes c ϵ {0, …, k-1} using k-1 parameter vectors β0, …, βk-2 ϵ Rd , [11]. Allow the response variable Y ϵ {1, …, k} have k possible values/categories. A form of the MLR is given by

$$p(Y = r|x) = \frac{\exp\,(xT\,\beta r)}{\sum_{s=1}^{k} \exp\,(xT\,\beta s)} = \frac{\exp(\eta r)}{\sum_{s=1}^{k} \exp\,(\eta s)} \quad (1)$$

Where βTk = (βr0, …, βrp).

Here we have to choose a reference category, when category k is chosen, one sets βTk = (0, . . . , 0) producing ηk = 0. We can select any other category as a reference, mobile in this research. Fitting a model using a reference category k, the corresponding equation is

$$p(Y = r|x) = \frac{\exp\,(xT\ \beta r)}{1+\sum_{s=1}^{q}\exp\,(xT\ \beta s)} \quad for \quad r = 1, \dots, q) \tag{2}$$

for r=1,…,q

We focused on the coefficients and p-values to see the contribution of each variable under study on the probability of using each electronic banking channel. The coefficients show the statistical relationship between the variables under test and the dependent variables. P-value is the preselected level of significance. We consider any variable with p-value less than 0.05 as significant and has an important effect on the dependents variables in this study [12].

4 Data

In this study, a substantial amount of panel financial data was obtained from three Kuwaiti banks, which vary in assets size and banking types. More precisely we gathered the data from two conventional banks, medium and small, and one medium size Islamic bank. The data collected covered a span of ten years from 2008 to 2018. The data include retail customers only; number and value of their monthly online transactions in each online channel; automated teller machine, point of sale, online banking, mobile apps, automated call centre, specifying the type of transaction done through these channels; deposits, withdrawals, purchases, payments, and transfers. The data included customer characteristics and their product portfolios, specifically:

- Customer age
- Customer address (Governance within Kuwait)
- Customer gender
- Customer nationality
- Customer income on monthly basis
- Customer monthly transactions, including date of transaction, value of payment per transaction, transaction payment currency, type of card used, transaction type such as purchase, withdrawal, payments, deposits and transfers
- Number and balance of current account, saving account, retail deposit, and joint account
- Total amount, principle paid, interest paid, and outstanding balance of consumer loans and instalment loans
- Limit and outstanding balance of credit cards
- Date of the relationship with the bank

The extensive data collection from the three banks had some limitation, that each bank introduced their electronic channel in different periods. Previously, older systems registered two or more services under the same name.

Bank terminology changes may affect data analysis. For example, prior to 2017, bank call centres registered payments and transfers under the terminology "payments." From 2017 onward, banks started registering each service separately. For better comparison between channels, some of the older incomplete data was neglected to maintain consistency. It was decided to only use the electronic channels that are offering the same services in order to generate better comparisons. These channels are (1) call centre (2) internet and (3) mobile. These platforms provide payments and transfer services while the other channels provide different types of services like deposits and withdrawals.

Table 1 summarises the data for each one of the Banks. Table 2 shows the size of data set, data sample and code execution times for each bank dataset. Bank three final observations' size (given all parameters used for this bank) is larger compared to that of Bank one, which resulted in a bigger size of regression data set. From Table 2 we can see that the code execution time for Bank three is less than that of the other two banks despite the larger amount of regression data set. This is attributed to smaller size data needing more iterations in model estimation.

Table 1. Bank data

	Bank 1	Bank 2	Bank 3
Total customers under observation	295,801 customers	95,429 customers	392,995 customers
Time period	11 years	11 years	9 years
Number of transactions	Call centre: 209,454 Internet: 385,887 Mobile: 341,601	Call centre: 144,579 Internet: 278,010 Mobile: 9,384	Call centre: 34,118 Internet: 1,244,746 Mobile: 3,821,546
Final sample size (given all parameters used for this bank)	7,083 customers	3,014 customers	17,743 customers
Final observations size (given all parameters used for this bank)	108,071 observations	54,494 observations	363,457 observations
Banks type	Conventional	Conventional	Islamic

Table 2. Size of data set & Code execution time

Bank	Size of whole data set (GBs)	Cleaning the data (hrs)	Size of regression data set	Fit MLR model (secs)	Predict probabilities for MLR (sec)
Bank 1	1.34	70	1.25 MB	182	33
Bank 2	0.05	7	0.62 MB	184	20
Bank 3	1.15	57	4.34 MB	66	75

5 Results

Tables 3 and 4 display the results from the multinomial logistic regression. The results show that as customers get older in age, they tend to prefer using call centres as their primary method of communication.

Table 3. Coefficients-relative log odds

Coefficients-relative log odds	**Bank 1**		**Bank 2**		**Bank 3**	
	Internet/ Mobile	Call centre/ Mobile	Internet/ Mobile	Call centre/ Mobile	Internet/ Mobile	Call centre/ Mobile
Intercept	-0.154	-0.723	-3.01	-0.393	-2.255	-7.016
x1-age	0.025	0.033	0.016	0.024	0.01	0.039
x2-gender	-0.156	-0.457	-0.496	-0.106	0.363	-0.915
x3-CA Count	-0.971	-1.118	-0.128	-1.061	-0.131	-0.258
x4-Da Count	0.623	0.505	-0.856	-2.383	0.039	0.055
x5-JA Count	0.206	0.374	-1.152	0.905	N/A	N/A
x6-SA Count	-0.359	-0.388	-1.889	-2.151	-0.247	-0.276
x7-Continent num B1/nationality B2	0.023	0.151	-0.003	-0.01	N/A	N/A
x8-CL Count	0.119	0.186	-0.001	-2.552	-0.037	0.211
x9-Income	4.38E-05	3.58E-05	-4.15E-06	-1.20E-03	-1.12E-05	-8.52E-05
x10-Ins.L Count	-0.195	0.023	0.506	-0.371	-0.042	0.375
x11-Credit card lim.	0	-6.84E-05	0	2.14E-04	-2.76E-05	-2.34E-05
x12-Credit card out.	0	1.81E-05	0	-4.45E-04	-1.09E-05	-1.51E-04
X13-Relationship duration in Y	N/A	N/A	0.043	0.264	0.057	0.244

Table 4. P-Value

P-value	**Bank 1**		**Bank 2**		**Bank 3**	
	Internet/ Mobile	Call centre/ Mobile	Internet/ Mobile	Call centre/ Mobile	Internet/Mobile	Call centre/ Mobile
Intercept	2.27E-05	1.22E-85	2.20E-74	4.89E-03	0.00E+00	0.00E+00
x1-age	2.84E-291	0.00E+00	5.51799E-07	3.26299E-24	3.41509180109000e-312	0.00E+00
x2-gender	2.00E-26	2.64E-209	2.92E-20	1.29E-02	0.00E+00	0.00E+00
x3-CA Count	0	0	0.003367069	4.0863E-167	6.2748E-105	1.41795E-31
x4-Da Count	1.63E-265	1.21E-142	1.93E-02	9.67E-14	2.84E-35	2.82E-07
x5-JA Count	0.72240804	0.519496747	8.41607E-20	6.99E-45	N/A	N/A
x6-SA Count	0	0	4.42419E-50	1.66E-84	0.00E+00	1.86E-218
x7-Continent num B1/nationality B2	0.00880377	4.81E-73	2.40E-06	8.72E-45	N/A	N/A
x8-CL Count	1.46E-81	8.83E-197	9.80E-01	0.00E+00	3.29E-19	2.29E-42
x9-Income	3.16E-110	9.32E-62	7.58E-01	0.00E+00	6.64E-27	1.17E-92
x10-Ins.L Count	0	3.34E-11	9.89E-81	7.01E-34	5.24E-55	0.00E+00
x11-Credit card lim.	0	1.27E-56	8.52E-31	5.27E-10	3.10E-130	1.19E-07
x12-Credit card out.	2.00E-152	0.000173234	3.04397E-92	4.09E-80	2.31E-05	4.09E-48
X13-Relationship duration in Y	N/A	N/A	1.85198E-10	0.00E+00	0.00E+00	0

Figure 1 explicitly shows that the second most favoured method of communication when people gets older is the internet, while mobile usage is their least favourable out of the three platforms. Results also indicate that the longer a customer's relationship with the bank is, the higher their probability of using call centre and internet than mobile is.

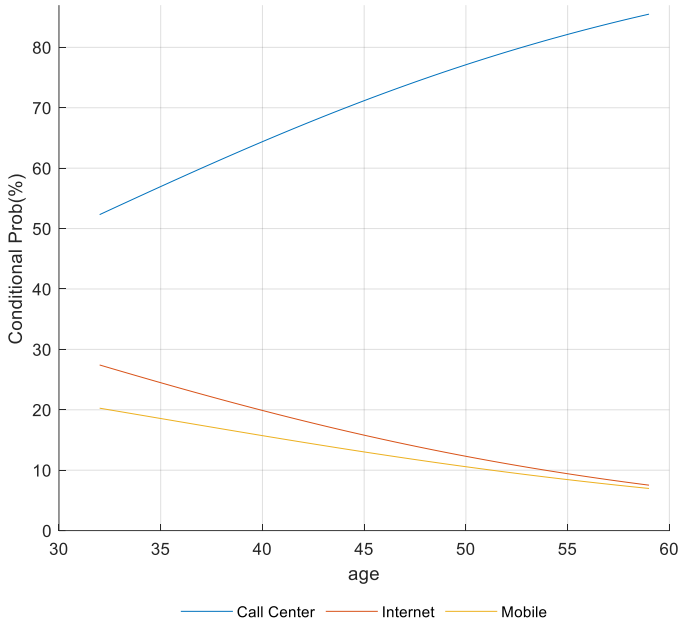


Fig. 1. Probabilities of electronic banking usage considering customer age

It was also found that a higher probability of males compared to females exists in terms of usage of mobile, rather than internet and call centres, given that the other variables remain the same. In Bank 3, the probability increased by 1.43 times with males compared to females using the internet over mobile and call centres, given everything else remain the same.

Regarding the four types of banking accounts, current, saving, deposit, and joint, the outcomes show that as the number of accounts increased, the probability of customers using mobile banking followed by the internet banking also increased. A noticeable decrease in call centre usage was observed as well. Data indicates customers with more than one deposit account use the internet and call centre, followed closely by mobile. The same data shows a noticeable decreasing trend in call centre usage and an increasing trend in internet and mobile usage. It should also be noted that the greater the number of joint accounts owned by a customer, the higher the probability of that customer using a call centre than other channels. Again, this result is aligned with the decreasing trend in call centre and increasing trend in mobile. Bank 1 showed insignificant relation of this variable given the high p-value.

The analysis also included the number of loans, consumer or instalment, held by each customer and whether this influenced the use of electronic channels of banking. We found that in Bank 1 and Bank 3, the greater the number of consumer loans per customer, the more that each individual customer tend to use call centre when compared to two other electronic channels. However, it also showed a shift from call centre to mobile in recent years. Bank 2 results display that the greater the number of consumer loans per customer, the greater the usage of mobile followed by call centre. Figure 2

results in Bank 2 also showed that there is insignificant impact of this variable on the customer decision of whether to use internet or mobile banking.

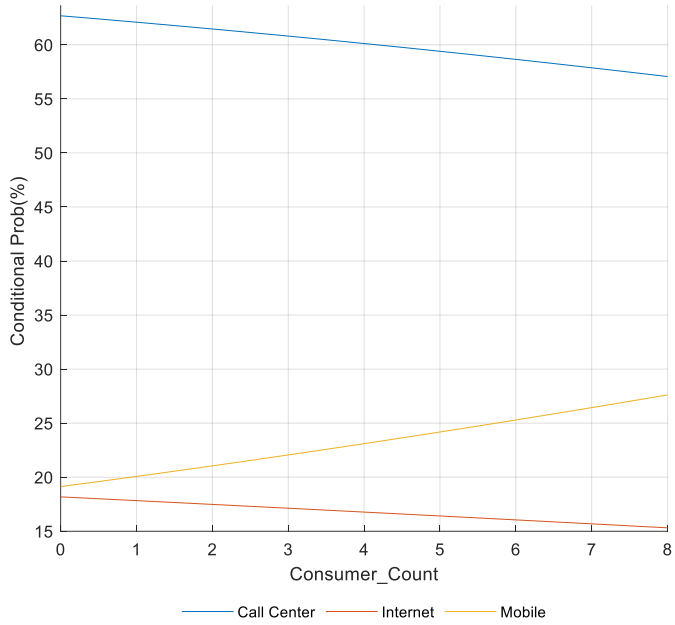


Fig. 2. Probabilities of electronic banking usage considering the number of consumer loans owned by customers.

Regarding the instalment loans, Bank 1 and Bank 3 showed that the greater the number of instalment loans per customer, the more that customer tend to use call centre when compared to the other electronic channels. Data also showed a shift from call centre to mobile in recent years. Bank 2 results in figure 3 state that the higher the number of instalment loans per customer, the higher the use of internet then mobile followed by call centre.

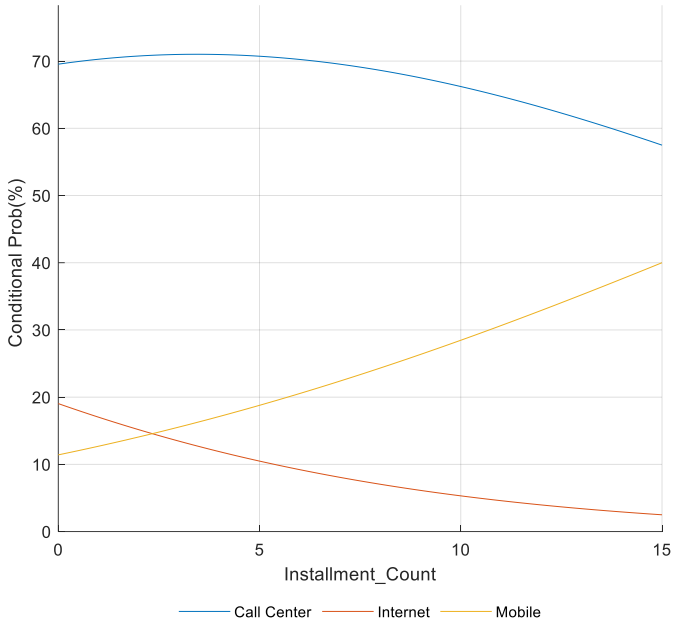


Fig. 3. Probabilities of electronic banking usage considering the number of instalment loans owned by customers.

Results show a weak correlation between customer income level and their use of electronic banking. The high p-value in the results denotes that income level cannot determine customer preferences between internet and mobile. Two variables were also tested: credit card limit and credit card outstanding balance. In all three banks, there is a very small coefficient for those variables which indicates the weak correlation of this variables with the preference of usage of electronic banking channels. Another variable, continent in Bank 1 and nationality in Bank 2, had no significant effect on customers behaviour in using the electronic channels.

6 Conclusion

The purpose of this research was to calculate the comparison probability usage of each banking channel for every client based on their information. The literature review discovered issues that were not addressed before; these are discussed in the findings section. Now, using the multinomial logistic regression, a prediction formula was created so that each bank can predict which banking channel will be used most frequently and effectively by each of their customers.

This prediction formula has capability to help banks allocate the correct marketing tools to the appropriate customer. One of the main potential impacts of this research is to create a better understanding of customer needs by offering quality products rather

than quantity. By precisely offering the most suitable products and services, customer engagement will be more productive and trust between banks and customers will increase. The identified prediction model will allow banks to reduce unnecessary marketing and capital expenditure, reduce the number of employees and ultimately increase their profit. By predicting customers' preference between available banking channels, banks will directly advertise each specific channel to the client interested in that channel, so banks can allocate the correct marketing tools in the right place. Realizing this information in advance will help banks to reduce their capital expenditure by knowing exactly where to invest their capital. For example, if banks recognized that most customers are shifting to mobile application, banks will alter their decision by enhancing that channel, mobile application, instead of putting additional amount of cost on all other channels. By knowing customers preferences, banks can reduce unnecessary number of employees in some channels. On the other hand, when banks understand customers' need, this will help to increase customer satisfaction resulting in more productivity.

Most recent works, in their data collection, they rely on questionnaires on limited group of respondents while we used real statistical numerical data on larger population. They focused on customer characteristics like age, gender, level of education as well as their behavioural factors with willingness in taking risk and paying price of the advanced technology. In our study we include the three mentioned characteristics along with nationality and we also consider their financial product portfolio which was not addressed. The research also used statistical modelling to solve financial problem and its quite comparative to our MATLAB usage. E-banking predictors can send meaningful and timely information to customers.

In summary, mobile banking technology has a lot of benefits for clients and banks. We collected real time data of e-banking channel usage and identified the factors influence its usage. To affectively adopt this platform, we need to know and study all the variables that could affect customer behaviour and increase their satisfaction level. In The Long Run, banks will gain more numbers of loyal customers.

References

1. Sadiq Sohail, M., Shanmugham, B.: E-banking and customer preferences in Malaysia: An empirical investigation. Inf. Sci. (Ny). 150, 207–217 (2003). https://doi.org/https://doi.org/10.1016/S0020-0255(02)00378-X.
2. Shankar, A., Jebarajakirthy, C., Ashaduzzaman, M.: How do electronic word of mouth practices contribute to mobile banking adoption? J. Retail. Consum. Serv. 52, 101920 (2020). https://doi.org/10.1016/j.jretconser.2019.101920.
3. Moazenzadeh, D., Hamidi, H.: Analysis and Development of Technology Acceptance Model in Mobile Bank Field. Int. J. Eng. 31, 1521–1528 (2018). https://doi.org/10.5829/ije.2018.31.09c.07.
4. Malaquias, R.F., Hwang, Y.: Mobile banking use: A comparative study with Brazilian and U.S. participants. Int. J. Inf. Manage. 44, 132–140 (2019).

https://doi.org/10.1016/j.ijinfomgt.2018.10.004.
5. Baabdullah, A.M., Alalwan, A.A., Rana, N.P., Kizgin, H., Patil, P.: Consumer use of mobile banking (M-Banking) in Saudi Arabia: Towards an integrated model. Int. J. Inf. Manage. 44, 38–52 (2019). https://doi.org/10.1016/j.ijinfomgt.2018.09.002.
6. Anouze, A.L.M., Alamro, A.S.: Factors affecting intention to use e-banking in Jordan. (2020). https://doi.org/10.1108/IJBM-10-2018-0271.
7. Choudrie, J., Junior, C.O., McKenna, B., Richter, S.: Understanding and conceptualising the adoption, use and diffusion of mobile banking in older adults: A research agenda and conceptual framework. J. Bus. Res. 88, 449–465 (2018). https://doi.org/10.1016/j.jbusres.2017.11.029.
8. McCullagh, P., Nelder FRS, J.A.: Generalized Linear Models. Chapman and Hall (1990).
9. The MathWorks Inc.: MATLAB. R2020a. (2020).
10. Okasha, M.K., Shehada, M.A.: CLASSIFICATION OF ANEMIC PALESTINIAN CHILDREN USING THE MULTINOMIAL LOGISTIC REGRESSION MODEL. Mahmoud. Int. J. Adv. Res. 4, 560–573 (2016).
11. Carpenter, B.: Lazy Sparse Stochastic Gradient Descent for Regularized Mutlinomial Logistic Regression Maximum a Posteriori Estimate. Alias-i, Inc. Tech. Rep, 1–20 (2008).
12. Gibbons, J.D., Pratt, J.W.: and Methodology * P-values : Interpretation. 29, 829–834 (1975).



https://dspace.org/

Using client's characteristics and their financial products to predict their usage of banking electronic channels

Alsanousi, Hessa; Albarrak, Najla; Moulitsas, Irene; Filippone, Salvatore

2021-10-27
Springer
https://doi.org/10.1007/978-981-16-2102-4_15
Alsanousi H, Albarrak N, Moulitsas I, Filippone S. (2021) Using client's characteristics and their financial products to predict their usage of banking electronic channels. In: 6th International Congress on Information and Communication Technology, 25-26 February 2021, London, UK

Downloaded from a DSpace repository. https://dspace.org/