SETSCI Conference Proceedings, 4 (1): 180-183, 2019

ISAS 2019
3rd International Symposium on Innovative Approaches in Scientific Studies
Section: Engineering and Natural Sciences
April 19 – 21, 2019, Ankara, Turkey

A Data Analysis and Behavior Model for Study of Consumer Service in the Financial Sector

Sultan Tatlılıoğlu[1*+], Dionysis Goularas[1]

[1]Computer Engineering/Science Institute, Yeditepe University, Istanbul, Turkey
[*]Corresponding author: sultanta@gmail.com
[+]Speaker: sultanta@gmail.com
Presentation/Paper Type: Oral / Full Paper

Abstract – In recent years, the financial sector has been considerably developed, especially in electronic banking. As individual users in the financial sector use e-banking and e-finance systems, the concept of loyalty in the banking sector has shifted from traditional branch banking where staff and customers have personal relations to self-service online banking. As expected, online banking competition is based on pricing and service quality. In this work special pricing, discounts, campaigns data and potential customers' searches are analyzed from available data. Modeling and segmentation were carried out with data mining and machine learning methods.

Keywords – Data Mining, e-banking, e-finance, Machine Learning, Modeling, Segmentation.

I. Introduction

Modern customers are more knowledgeable, educated, sophisticated, independent, value and motivational-oriented than before. By conducting their own research and analysis themselves, they learn how to determine the alternatives in the market and the product that is most appropriate for them according to their own budget. Search and comparison engines like Google, YouTube video platform, web search, CPU benchmarks, Top10 reviews, social platforms are the most widely used access resources. With this way, customers find alternative products and easily compare them with few clicks.

This research is focused on events that will be repetitive and innovative in customer's personal financial decisions and payments. More precisely, we are focusing on payments such as taxes and fees to be made at certain times of the year. In order to cover the above payments, customers are searching for possible credit applications for personal needs such as Religious Feast (Bayram) expenses, Motor Vehicle Tax or Real Estate Tax. As there is no sense of loyalty or habit for people in this product, customers can make the payment by taking a credit from a bank of their choice.

II. Method

In this research we used data from Google Trends. Choi and Varian [1] have shown that data from Google Trends can be linked to current values of various economic indicators, including automobile sales, unemployment claims, travel destination planning and consumer confidence. A very recent study has shown that Internet users from countries with a higher per capita GDP are more likely to search for information about years in the future than years in the past [2].

A. Number of Researches with Google Trends

Figure 1 shows the number of studies that are using Google trends. Among them, some studies utilized Google Trends in order to investigate economic activity such as product purchasing behaviour in the market of both property and workforce [3], Google Trends is applied in the domain of medicine as well as in social science. Using monthly data, they report a strong relationship between search keywords unemployment rates [4]. Search query data are helpful for predicting quickly changing and complicated trends.

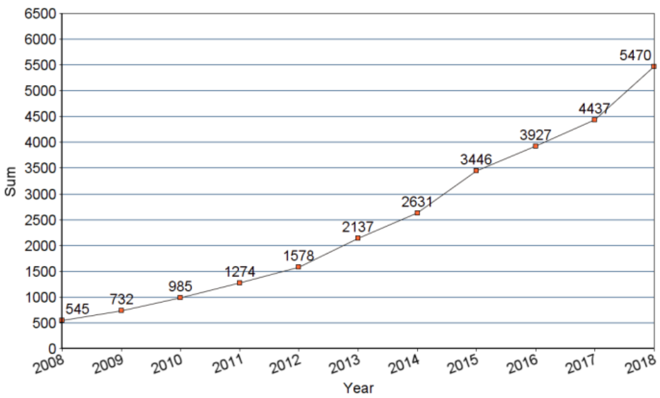

Fig. 1 Number of researches with Google Trends

Some other studies utilized Google Trends to investigate economic activity such as product purchasing behaviour in the market of both property and workforce [5]. Google Trends is also applied to investigate subjects in medicine as well as in social science. Askitas and Zimmermann suggest that Google Trends information are a suitable feedback for policy-making. Using monthly German data, they report a strong relationship between search keywords unemployment rates from 2004-2009 [6]. Search query data are helpful for predicting quickly changing and complicated trends. D'Amuri and Marcucci [7], reported that the outcomes of their search query analysis produce more accurate estimations than the state-level survey or expert predictions and they employed Google Trends to estimate the employment rate in the US. In France, by utilizing information from Google Trends of related queries, Scholars [8] estimated the unemployment rate. Doornik [9] and

Ginsberg et al. [10] utilize Google Trends query on influenza virus inspection. Cooper, Mallon, Leadbetter, Pollack, and Peipins [11] reported that the Google search query for particular cancers complements the expected occurrence from 2001 to 2003. A high association among the inquiries on the Google search for flu queries Canada during 2004 to 2005 season is reported by Eysenbach [12]. Polgreen et al. [13] reported that the frequency of notified case of influenza is associated with the number of Google inquiries for influenza-related queries during 2004-2008. Hulth, et al. [14] inform similar outcomes in a research of the search queries submitted to the medical network. The notification of suicide-related signs in public health information is too delayed to impact social factors as stated by McCarthy [15] and Gunn HI and Lester [16]. Using Google trend, Gunn HI and Lester [16] suggested that the suicide trend can be monitored faster in time than the official's suicide statistics report by analysing the correlation between Google Trends suicide queries and suicide rates. Sueki [17] reports that queries for "depression" could notify the public health officials to an upcoming increase the rate of suicide in Japan. In his work he examined the fluctuations in Google search index for depression and suicide. Moreover, these previous research suggested the benefit of utilizing Google Trends to obtain information on real-time by collecting and to prevail over errors which may be generated from overdue data publish.

B. Problem Domain: Financial Markets Campaigns

Financial markets are a prime target for such quantitative investigations [18],[19]. Movements in the markets exert immense impacts on personal fortunes and geopolitical events, generating considerable scientific attention to this subject [20],[21]. For example, a range of recent studies have focused on modeling financial markets [22],[23] and on performing network analysis [24],[25].

1) Religious Feast (Bayram) Credit Campaigns

Figure 2 presents the Google Trends search charts in financial data for the last 5 years (2014-2018) with feast and credit keywords. The calendar of Ramadan and Qurban feasts shifts 10 days each year .

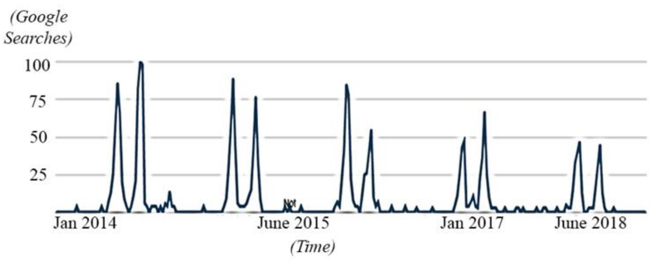

Fig. 2 Google Trends graphic of feast and credit keywords for the years 2014-2018.

2) Motor Vehicle Tax Campaigns

Figure 3 presents the Google Trends search chart in financial data of the last 5 years with MTV and instalment keywords. Motor Vehicle Tax times are between 1 – 31 January and 1 – 31 July.

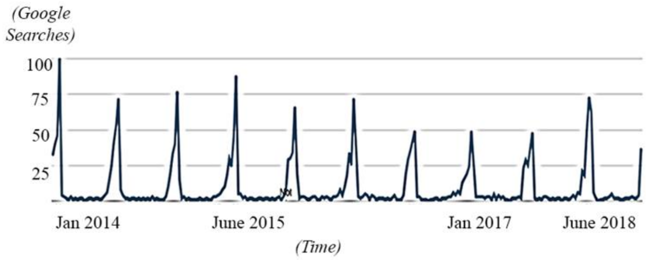

Fig. 3 Google Trends graphic of MTV and instalment keywords for the years 2014-2018

3) Real Estate Tax Campaigns

Figure 4 shows the Google Trends search chart in financial data of the last 5 years with real, estate, tax and instalment keywords. Real Estate Tax times are between 1 March – 31 May and 1 - 30 November.

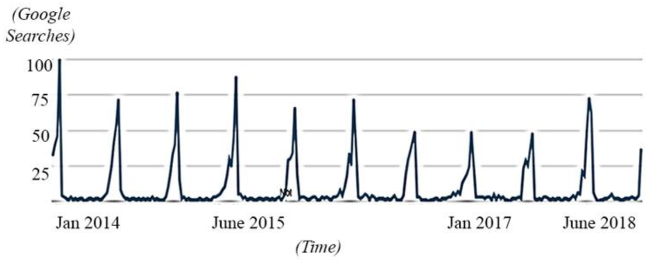

Fig. 4 Google Trends graphic of real, estate, tax and instalment keywords for the years 2014-2018

C. Customer Behaviour Analysis Data Fields

One widely known unsupervised classification algorithm based on clustering the data into local regions, is the k-means algorithm [26]. K-means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem [27],[28],[29]. K-means is an iterative procedure, including that of Legendre and Legendre [30] , to place cluster centers, which quickly converges to a local minimum of its objective function [31],[32].

III. RESULTS

In this section we are presenting the results of this study. The customers and bank campaigns data are from the year 2014. The following questions are meant to reveal the relation between the difference of the actual events, the financial campaigns and the customers searches habits based on the Google Trends.

A. How Many Days Are Customers Starting to Search before the Event Start Date?

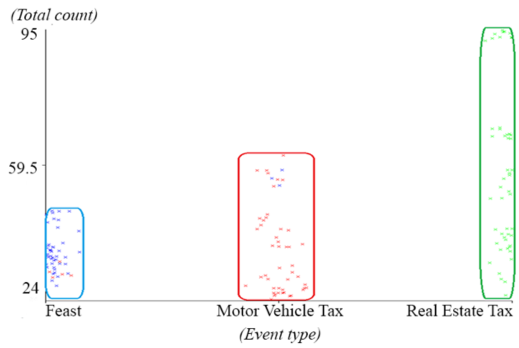

Fig. 5 K-means result days difference between the event start date and Google trends. Feast, Motor Vehicle Tax and Real Estate Tax are presented with blue, red and green colours correspondingly.

Figure 5 displays the difference in time between the event and the researches were made on the selected events. We observe that Real Estate Tax was searched first and Feast credit was the last searched event.

B. *How Many Days Before was the Peak Happen from the Event Start Date?*

Figure 6 presents the difference between the Peak event and its start date. Real Estate Tax and Motor Vehicle Tax searches are began then the peak being at the same time. Feast is spreading in a longer period. Since the amount of feast credit is bigger than others, the customers have a tendency to think for a longer period.

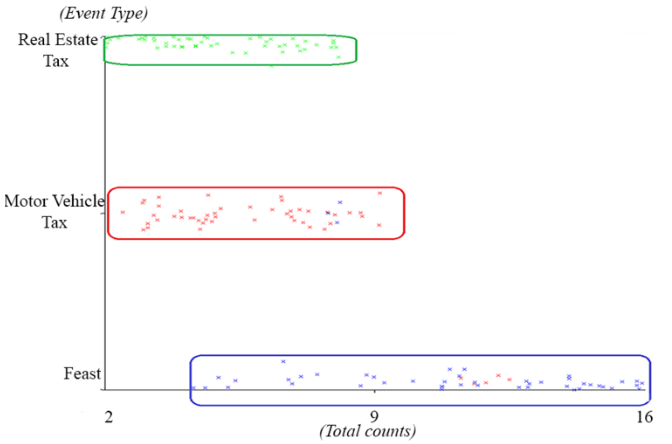

Fig. 6 K-means results on day differences between the Peak event and that event start date. Feast, Motor Vehicle Tax and Real Estate Tax are presented with blue, red and green colours correspondingly.

C. *How Many Days the Campaign Started Before the Event Start Date? (Event Start Date - Campaign Start Date)*

In figure 7, we can see the difference between the start of the campaign and the event start date. The feast campaigns started before the event. The latest campaign was the real estate tax.

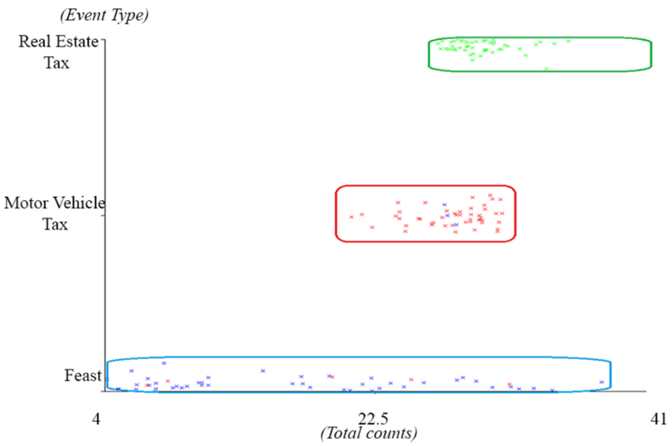

Fig. 7 K-means results on the difference between the start of the campaign and the event start date. Feast, Motor Vehicle Tax and Real Estate Tax are presented with blue, red and green colours correspondingly.

IV. DISCUSSION

In summary, our results are consistent with the hypothesis that during the period we investigate, Google Trends data did not only reflect aspects of the current state of the economy, but may have also provided some insight into future trends in the behaviour of economic actors.

The usage popularity of this keyword in searching activities also indicates information demand of banking services and products. This findings supports that the information search is one of the most important step in consumer decision making [33] as well as Milner and Rosenstreich financial behaviour model [34].

The results of our investigation suggest that combining large behavioural data sets such as financial trading data with data on search query volumes may open up new insights into different stages of large-scale collective decision-making.

V. CONCLUSION

In this study we presented an analysis of the consumers behaviour on three financial subjects based on google trends and the corresponding response of banks on these items. With this work, using real customer and financial institutions data we attempted to find answers to the following questions: How many instalments should the bank make in the campaign? Can we look at peak times in searches and set the suitable date and parameters such as interest ratio, discount or bonus for the campaign? The results shows that the financial institutions and the customers' search behaviour have a time difference between them and the reason for it and its consequences needs to be further investigated.

REFERENCES

[1] Choi, H. & Varian, H. Predicting the Present with Google Trends. The Economic Record 88, 2–9 (2012).

[2] Preis, T., Moat, H. S., Stanley, H. E. & Bishop, S. R. Quantifying the Advantage of Looking Forward. Scientific Reports 2, 350 (2012).

[3] P. Smith, "Google's MIDAS touch: Predicting UK unemployment with internet search data," J. Forecast. 35(3), pp. 263-284,2016.

[4] N. Askitas, K. F. Zimmermann, "Google econometrics and unemployment forecasting," Appl. Econ. Quart. 55(2), pp. 107-120,2009.

[5] P. Smith, "Google's MIDAS touch: Predicting UK unemployment with internet search data," J. Forecast. 35(3), pp. 263-284,2016.

[6] N. Askitas, K. F. Zimmermann, "Google econometrics and unemployment forecasting," Appl. Econ. Quart. 55(2), pp. 107-120,2009.

[7] F. D'Amuri, J. Marcucci, "Google it! Forecasting the US unemployment rate with a Google job search index," FEEM Working Paper 31,2010.

[8] Y. Fondeur, F. Karame, "Can Google data help predict French youth unemployment?" Econ. Model. 30, pp. 117-125,2013.

[9] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S.Smolinski, L. Brilliant, "Detecting influenza epidemics using search engine query data," Nat. 457(7232), pp. 1012-1014,2009.

[10] J. A. Doornik, "Improving the timeliness of data on influenza-like illnesses using Google search data," Working Paper, 2009= J. A. Doornik, "Improving the timeliness of data on influenza-like illnesses using Google search data," Working Paper, 2009.

[11] P. C. Cooper, P. K. Mallon, S. Leadbetter, A. L. Pollack, A. L.Peipins, "Cancer internet search activity on a major search engine, United States 2001-2003,"/. Med. Internet Res. 7(3), e36,2005.

[12] G. Eysenbach, "Infbdemiology: Tracking flu-related searches on the web for syndromic surveillance," AMIA Annual Symposium Proceedings, pp. 244-248,2006.

[13] P. M. Polgreen, Y. Chen, D. M. Pennock, F. D. Nelson, R. A. Weinstein, "Using internet searches for influenza surveillance," Clin. Infect. Dis. 47(11), pp.1443-1448,2008.

[14] A. Hulth, G. Rydevik, A. Linde, "Web queries as a source for syndromic surveillance," PLOS ONE 4(2) e4378,2009.

[15] M. J. McCarthy, "Internet monitoring of suicide risk in the population," J. Affect. Disord. 122(3), pp. 277-279,2010.
[16] J. F. Gunn HI, D. Lester, "Using Google searches on the internet to monitor suicidal behavior," J. Affect. Disord. 148 (2-3), pp.411-412,2013.
[17] H. Sueki, "Does the volume of Internet searches using suicide related search terms influence the suicide death rate: Data from 2004 to 2009 in Japan," Psychiatry Clin. Neurosci. 65(4), pp.392-394,2011.
[18] Fehr, E. Behavioral science - The economics of impatience. Nature 415, 269–272(2002).
[19] Shleifer, A. Inefficient Markets: An Introduction to Behavioral Finance (Oxford University Press, Oxford, 2000).
[20] Lillo, F., Farmer, J. D. & Mantegna, R. N. Econophysics - Master curve for price impact function. Nature 421, 129–130 (2003).
[21] Watanabe, K., Takayasu, H. & Takayasu, M. A mathematical definition of the financial bubbles and crashes. Physica A 383, 120–124(2007).
[22] Bouchaud, J. P., Matacz, A. & Potters, M. Leverage effect in financial markets: the retarded volatility model. Physical Review Letters 87, 228701 (2001).
[23] Sornette, D. & von der Becke, S. Complexity clouds finance-risk models. Nature 471, 166 (2011).
[24] Schweitzer, F. et al. Economic Networks: The New Challenges. Science 325,422–425 (2009).
[25] Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. Nature 464, 1025–1028 (2010).
[26] J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations", Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, University of California Press, 1 pp.281-297, 1967.
[27] Kaufman, L. & Rousseeuw, P.J., Finding Groups in Data: An Introduction to Cluster Analysis, John Wiley and Sons, Inc., New York ,1990.
[28] Legendre, P. & L. Legendre, Numerical Ecology, Second edition. Elsevier Science BV, Amsterdam, the Netherlands, 1998.
[29] Bradley & Fayyad , "Scaling Clustering Algorithms to Large Databases", Proceedings of the Fifteenth International Conference on Machine Learning ICML98, pp.91-99, 1998.
[30] T. Kanungo,DM Mount, N. Netanyahu, Christine D. Piatko, Ruth Silverman & Angela Y. Wu, " An efficient kmeans clustering algorithm: Analysis and implementation", IEEE Trans. Pattern Analysis and Machine Intelligence, 24, pp.881 -892,2002.
[31] K. Krishna & Raghu Krishnapuram, " A Clustering Algorithm for Asymmetrically Related Data with Applications to Text Mining", Proceedings of the International Conference on Information and Knowledge Management(CIKM2001), Atlanta, Georgia, USA, pp.571-573, 2001.
[32] S. Guha, R. Rastogi & K. Shim, "CURE: An efficient clustering algorithm for large databases", Proceedings of ACM SIGMOD International Conference on Management of Data. New York, pp. 73-84, 1998.
[33] P. Kotler and K. L. Keller, "Marketing Management, Global Edition. 15th ed." Edinburgh Gate, England: Pearson Education Limited, pp.194-209, 2015.
[34] T. Milner and D. Rosenstreich, "A review of consumer decision making models and development of a new model for financial services." Journal of Financial Services Marketing, vol. 18(2), pp.106-120, 2013.