3-7-2023

Impact of Big Data Analytics on Banking: A Case Study

Wu He

Old Dominion University

Jui-Long Hung

Boise State University, andyhung@boisestate.edu

Lixin Liu

Yangzhou University

Publication Information

He, Wu; Hung, Jui-Long; and Liu, Lixin. (2023). "Impact of Big Data Analytics on Banking: A Case Study". *Journal of Enterprise Information Management, 36*(2), 459-479. https://doi.org/10.1108/JEIM-05-2020-0176

Impact of Big Data Analytics on Banking: A Case Study

Wu He
Old Dominion University

Jui-Long Hung
Boise State University

Lixin Liu
Yangzhou University

Abstract

Purpose

The paper aims to help enterprises gain valuable knowledge about big data implementation in practice and improve their information management ability, as they accumulate experience, to reuse or adapt the proposed method to achieve a sustainable competitive advantage.

Design/Methodology/Approach

Guided by the theory of Technological Frames of Reference (TFR) and Transaction Cost Theory (TCT), this paper describes a real-world case study in the banking industry to explain how to help enterprises leverage big data analytics for changes. Through close integration with bank's daily operations and strategic planning, the case study shows how the analytics team frame the challenge and analyze the data with two analytic models - customer segmentation (unsupervised) and product affinity prediction (supervised), to initiate the adoption of big data analytics in precise marketing.

Findings

The study reported relevant findings from a longitudinal data analysis and identified some key success factors. First, non-technical factors, for example intuitive analytics results, appropriate evaluation baseline, multiple-wave implementation, and selection of marketing channels critically influence big data implementation progress in organizations. Second, a successful campaign also relies on technical factors. For example, the clustering analytics could promote customers' response rates, and the product affinity prediction model could boost efficient transaction and lower time costs.

Originality/Value

For theoretical contribution, this paper verified that the outstanding characteristics of online mutual fund platforms brought up by Nagle, Seamans, & Tadelis (2010) could not guarantee organizations' competitive advantages from the aspect of TCT.

Keyword: transaction cost theory, big data analytics, enterprise information management, banking industry, precise marketing

1. Introduction

Big data analytics has received increasing attention in organizations especially financial institutions in the last few years. As the set of techniques that is being used to discover hidden knowledge and customer value, big data analytics has the potential to enable financial institutions to reap many benefits including increased earnings and reduced fraud losses (Naveira et al., 2018). However, there are challenges to successful use of big data analytics in the banking industry, including cost consideration and lack of required human resources and skills (Raguseo, 2018). In addition, traditional banks face competitions from highly digitalized companies, such as online mutual fund and online insurance platforms. These companies have adopted big data analytics heavily for marketing. Many banks have yet to realize the full potential of leveraging big data analytics to improve returns from their decision processes and business operations (Naveira et al., 2018). Mikalef et al. (2019) reported that managerial skills are critical for gaining values from big data analytics as data resources and technical skills are. Currently, many banks are not familiar with successful practices in terms of how to successfully apply big data analytics deeply into their culture, decision processes, and business operations.

This paper aims to help enterprises better leverage big data analytics for improving information management, operation and decision making. After all, the value of big data analytics depends on whether the insight generated by big data analytics can help improve firm operation or performance. As some personal or organizational factors may obstruct effective implementation of big data analytics in organizations, it is better to follow well-defined methodology or practices for implementation (Mikalef et al., 2019). Similar to other published case studies in organizations or governments such as Dolci et al. (2014) and Jones (2012), this paper discusses a real-world case study in an Asian bank – from planning to implementation – and observes related impacts over a longer period. The overarching research question for our study is: *How can big data analytics be effectively adopted to help banks improve performance?*

The remainder of the paper proceeds as follows. The second section provides a brief literature review about the Technological Frames of Reference (TFR), Transaction Cost Theory (TCT), adoption of big data analytics, and characteristics of big data analytics in the banking industry. Section 3 describes a case study and the related findings. Section 4 discusses the theoretical and practical implications. Section 5 concludes the paper.

2. Literature Review

2.1. Theoretical Foundations

Orlikowski and Gash (1994) propose TFR as a theoretic approach to study interpretation related to IT in organizations. The approach defined technological frames as the knowledge and expectations that guide actors' interpretations and actions related to IT. They further stated that social groups have shared frames and that differences in these groups' frames can inhibit effective deployment of a technology (Davidson, 2006). Organization members make sense of information technologies in different ways, which could influence their actions related to IT. Thus, it could be ineffective for achieving desired benefits when various groups in the organization have incongruent frames and interpret technology and related communication differently. Prior studies indicate that incongruence between frames could cause issues in technology implementation projects (Orlikowski & Gash 1994) as well as skepticism toward the adoption of new and better technology (Davis & Hufnagel 2007).

TCT (Williamson, 1975, 1985) suggests that the optimum organizational structure is the one that achieves economic efficiency by minimizing the costs of exchange. Williamson defined transaction costs broadly as the costs of running the economic system of companies. Thus, cost is the primary determinant of a decision. Transaction cost theorists suggest that the total cost acquired by a firm can be grouped into two major components: (1) transaction costs and (2) production costs. Transaction costs, often known as coordination costs, are defined as the costs of "*all the information processing necessary to coordinate the work of people and machines that perform the primary processes*", whereas production costs include the costs incurred from "*the physical or other primary processes necessary to create and distribute the goods or services being produced*" (Klimis, 2007). When making 'make-or-buy' decisions, decision makers must weigh up the production and transaction costs associated with implementing a transaction, such as in-house production versus outsourcing (Aubert & Weber, 2001).

Williamson (1985) discussed factors which might increase transaction costs. The factors include bounded rationality, opportunism, uncertainty and complexity, small number transaction, information asymmetry, and atmosphere. In addition, asset specificity, uncertainty, and frequency might also influence transaction cost. Heide (1994) defined asset specificity as the investments in physical or human assets that are devoted to a business partner and whose redeployment involves considerable switching costs. Uncertainty refers to the level of uncertainty associated with a transaction which especially associates with human's inability in solving complex problems. Transaction frequency refers to the frequency with which transactions recur (Williamson, 1985).

2.2. The Adoption of Big Data Analytics

While many companies realize that adopting big data analytics for decision making or value creation is beneficial (Shirazi & Mohammadi, 2019), not all industries and organizations had implemented successful big data projects. Raguseo (2018) noted several concerns, such as security and privacy issues, problems in managing large data size, lack of expertise and experience, and organizational resistance, have hindered many companies' adoption of big data analytics. Sanader & Marko (2017) believe that the use of big data analytics is critical in the banking industry, but banks face many non-technical challenges.

First, two general types of administrators or staffs can be found in the data analytics adoption. One tends to believe "personal big data rules" (derived from personal experiences) rather than machine-generated rules (Alles, 2015). The other regards "big data analytics" as an omnipotent pill with many unrealistic impressions (Michael & Lupton, 2016). The alleviation of possible negative effects from these types of administrators/staffs and ways to recognize the marketing organization are also key success factors.

Second, a successful technology adoption implementation involves consensus from top to bottom, along with the design of business processes, and the offering of appropriate incentives. Relying on experienced data scientists alone cannot ensure high profitable outcomes. Prior studies often focus on the technology aspect of data analytics (Elzamly et al., 2016; Srinivasan & Kamalakannan, 2018), there is not much discussion about ways in which a successful implementation of big data analytics can be achieved.

Third, big data methods still face some challenges to deal with a massive dataset with high dimensionality and to generate useful results without the appropriate variable selection process of dimension reduction (Sivarajah et al., 2017). As sophisticated analytic methods become more popular, the generated results can be difficult for people without professional training to understand. Therefore, the generation of easy-to-interpret results is crucial for successful decision-making support (De Laat, 2018).

Furthermore, Baker (2012) suggests that the technological, organizational, and environmental contexts are key successful factors for organizations. However, most studies have only outlined the adoption principles. Organizations who are new to big data analytics can find it hard to know where or how to begin. In the banking industry, two highly relevant areas related to big data analytics are precise marketing and risk management (Chedrawi et al., 2020). Compared with risk management which has been the core of the banking industry, precision marketing is relatively new to the banking industry.

Since big data analytics is complicated and often involves multiple business units, a feasible idea is to start with a smaller big data analytics project by identifying specific, high-value opportunities with big return. Sivarajah et al. (2017) suggested implementing big data analysis in sequence.

2.3. Characteristics of Big Data Analytics in the Banking Industry

Banks store a lot of business, transaction, and customer data. The quality of data in banking is better than many industries, because customers are required to provide correct and detailed personal information. A bank tracks every transaction of each customer's accounts, along with the customer's credit history, credit card transactions, and behaviors performed in online and in-person channels. Therefore, the challenge for the banking industry is the issue of high-dimensionality, rather than the high volume of data flow in e-commerce (Ngai, Xiu, & Chau, 2009).

Customer clustering is a popular analytics method used in the banking industry (Hendalianpour et al, 2017). The high dimension issue can make the analysis and interpretation operationally difficult (Marshall, Tang, & Milne, 2010). Ma, Baer, & Chakraborty (2015) noted that customer clustering can be a challenge when the sample size is large with assorted data issues. Their experiment included 30,000 customers but ended up with a higher density of observations near the global center of population; it classified about 50% of the customers into a single cluster. In addition, although there are different clustering methods for high-dimensional data (Esmin, Coelho, & Matwin, 2015), the analytics results might be still challenging for end-users to fully understand and interpret the results. Without easy-to-understand and meaningful analytics results, it will be difficult for data scientists to promote the analytics model and maximize the outcomes.

It is not easy to correctly evaluate the effects of big data analytics. A/B testing and baseline comparison are often used for big data analytics measurement. However, since operations are ongoing, it can be hard to define an appropriate baseline. Revenue generation and response rate are two common indicators for performance measurement (Maiste, 2018). However, a high response rate is not necessarily equal to a high monetary return.

2.4. A Proposed Framework for Big Data Analytics in Organizations

In the banking industry, the IT department is a crucial business unit for maintaining daily operations, generating managerial reports for decision making, and evaluating new technologies. However, the IT department's contributions are often underestimated because their contributions are hard to be quantified in monetary numbers. Therefore, many IT department hope to highlight their contributions by adopting the big data analytics. The operations department also hopes to identify new sales opportunities. The finance department primarily expects to reduce risks by reducing the scale of implementation. The management aims to increase revenues and positive visibility in news or media. Therefore, a big data analytics project is often shaped by different parties, making it challenging to achieve the desired outcomes and benefits. Based on the aforementioned theoretical foundations, we came up with the following framework to guide and analyze the case study:

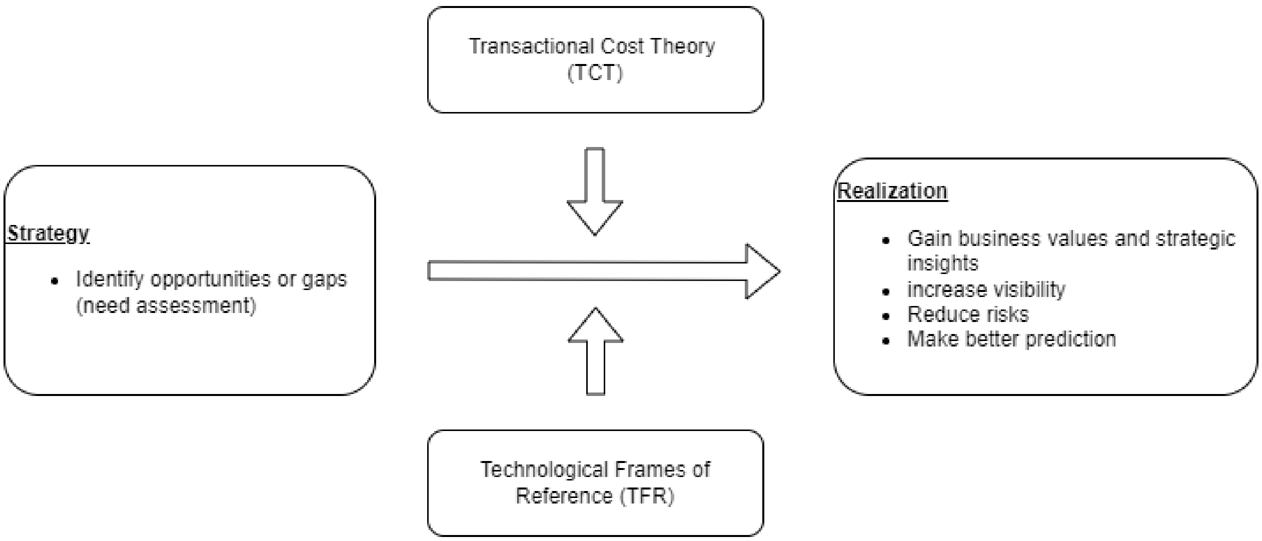

Figure 1. A framework to guide the implementation of big data analytics in organizations

Overall, organizations will need to do needs assessment first to identify needs, opportunities, or gaps. If the answer is positive, organizations will create roadmap for implementing big data analytics. The desired outcome is realization of various benefits including gaining business values (e.g. more earnings) and strategic insights for different managerial levels, increasing visibility, reducing risks, and making better predication. Studies show that numerous technology projects failed due to various personal, management, organizational and other contextual factors (Yeoh & Koronios, 2010). Thus, the TCT will be used to analyze individual tractions or activities with the aim to optimize the governance or organizational structure to achieve economic efficiency by minimizing the costs involved. The TFR will be used to address conflicts and skepticism among various groups in the organization on the implementation and adoption.

3. Case Study

The case study describes how a commercial bank in Taiwan adopted big data analytics in Customer Relationship Management to achieve their mission and goals. Using the proposed framework in figure 1, we framed the case study into the need assessment, the analytics process, and the management process.

3.1. Needs Assessment

In this case study, we tried to deeply understand the challenges of a top-five bank (referred to as "A-bank", hereafter) in Asia with a population of about twenty-three million. This bank is ranked as one of the top 250 banks worldwide with more than 100 branches, 30-plus overseas branches/representative offices, and over 6,000 domestic employees. Overall, A-bank's strength is in Corporate Finance, especially in Small and Medium-sized Enterprise (SME) loans. Compared to other top-tier commercial banks, A-bank's personal finance business is weaker. In 2015, A-bank set up its division of digital banking, through organization reengineering, to declare its ambitions in the field of FinTech.

A workshop was led by an external consulting team to initiate the cognitive process which involves the top-tier administrators and administrators of business units. The cognitive process is designed to identify potential weaknesses of personal finance for revenue improvement. Contextual analyses were conducted to facilitate related discussions. Two contextual challenges were well discussed during the workshop. First, in the past, the wealth management advisors across all branches would oversell the same group of high-end customers. This strategy meant that 90% of the revenues in the personal finance sector were contributed by only 10% of customers; this is significantly higher than the 80/20 principle. Second, as Figure 2 shows, there is a relationship between the average numbers of product holding and the bank customers' average ages. The results indicate a positive relationship between these two variables. However, the average number of products held reaches its height when customers are in the range of 51-60 years old (the average number of products held is 1.74). After the peak, the number drops significantly, since most of the customers are retired or are getting ready to retire after 60 years old. If those customers are on the receiving end of the same marketing strategy that is used to oversell the high-end customers, they might find it annoying. In addition, the revenue basis will outflow, as these customers age further. Below summarizes the strategies generated from the workshop.

The top-tier administrators adopted external consultants' suggestion to initiate a big data analytics project and provide necessary managerial efforts for successful implementation. The corresponding strategies include:

- All marketing campaigns should be shifted from the product-centered to customer-centered. Therefore, the analytics results should recommend products based on personal preferences.
- The analytics results can provide an overview of all customers for developing personalized CRM strategies and tracking their effectiveness.
- The analytics results can be applied to all product lines of personal finance.
- The analytics results can identify new potential customers without overselling the same groups of customers.

After numerous meetings, the external consultant proposed two solutions: 1) construct the customer segmentation to support strategic planning and decision makings in CRM; 2) construct the product affinity model for a personalized product recommendation. These two solutions not only aim to provide personalized financial services in personal banking, but also aim to increase the revenue base by stimulating inactive or by adding new customers.

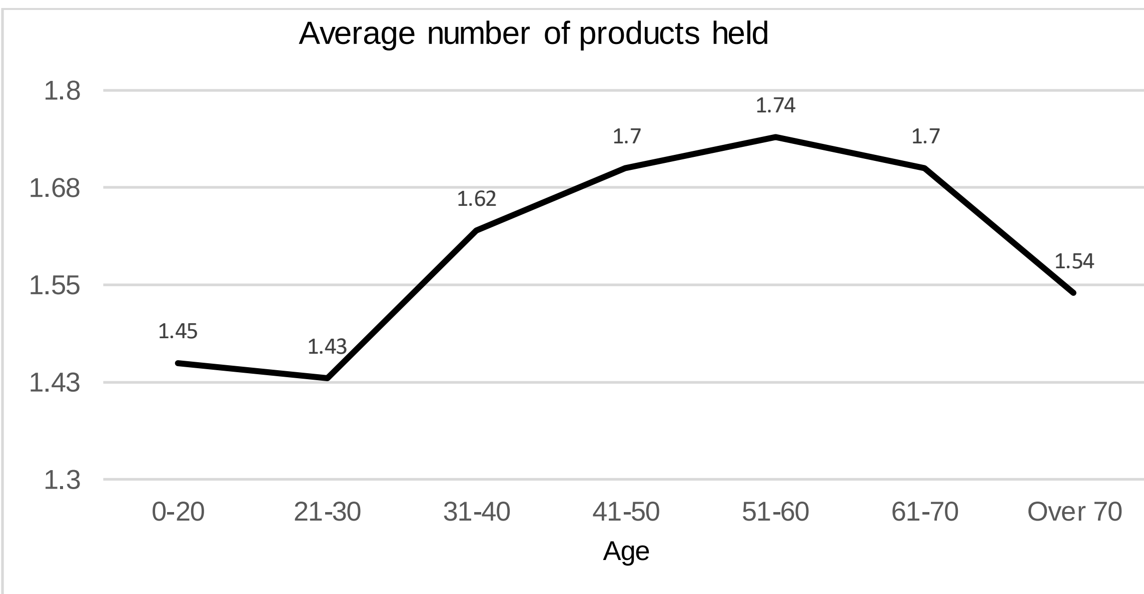

Figure 2. Relationship between age and product holding

3.2. Analytics Process

3.2.1. Customer Clustering Analysis

The customer clustering analysis is usually the first CRM model used in the banking industry. It provides an overview of all customers by classifying customers based on attribute similarities (Collica, 2017). However, it is hard to obtain good clustering results when you have over 2.5 million customers and over 700 potential (the long list) variables for a clustering analysis. A clustering analysis with a large sample size and many variables will encounter two issues. The first is the curse of dimensionality (Keogh & Mueen, 2017), which is a common challenge in the field of big data analytics, and the second is that there is often a high density of observations near the global mean (Khandare & Alvi, 2016). To solve these problems, as well as to make the data easier for bank staffs to interpret, the external consultant proposed a two-stage clustering approach that aimed to generate clusters that are both useful and meaningful.

The first stage is called strategic clustering. It aimed to generate clusters at the macro-level for strategic planning. In practice, whether the clustering model can generate stable and reproducible results is crucial for successful implementation (Melchiotti et al., 2017). The average silhouette was adopted as the stability indicator for model selection (Dudek, 2020). In addition, the final model was selected with two more criteria: (1) major characteristics of individual clusters should be intuitive to users and highly connected with financial products. and (2) the administrators can easily evaluate marketing efforts every six months by observing customer movements among clusters. Because of the above considerations, Customers' Assets Under Management (AUM) and Customer's Contributions were selected for the strategic clustering. AUM represents a customer's all product holdings of personal banking in monetary values. Customer's Contributions represent the summary of contributions from a customer's personal banking products (Haenlein, Kaplan, &Beeser, 2007). The combination of these two variables were intuitive to bankers. More important, because individual personal banking products has its own unique characteristics in terms of AUM and contributions, conducting clustering analysis with these two dimensions can generate clusters with distinctive product holding portfolios. After clusters were generated, additional fifty variables in customer profiles, product holdings, product account activities, Know Your Customer (KYC), and channel preferences were selected to identify unique characteristics of individual clusters for result interpretation. Other variables on the long list were excluded due to either different levels of collinearity or being redundant in the result interpretation (Sambandam, 2003).

Because the size of each cluster is still too large for the actual campaign activities, the results are suitable for strategic planning and CRM impact observations. The second stage is called the operational clustering analysis. The analysis further classified Stage 1 clusters into smaller clusters for use in actual marketing or CRM activities. After the first round of clustering analysis, the size of the individual clusters was about 20 thousand to 50 thousand customers. This size was manageable and could be further classified into smaller clusters by incorporating more variables.

The first stage of the cluster analysis incorporated two clustering and 50 interpretation variables. The second stage incorporated 20 clustering variables in product holdings and account activities and 30 interpretation variables in customer profiles, KYC, and channel preferences. Unique characteristics of individual clusters were identified via techniques of cluster profiling (Seth, 2021). The process mainly identified variables that differentiate individual clusters from other clusters. The cluster labels of the target cluster (the one to be profiled) and the rest of clusters were recoded into 1 and 0 respectively. Key variables that best differentiate the target cluster from others were ranked by feature importance generated by Random Forest algorithm (Alghofaili, 2020). In addition to key differentiation variables, basic statistics, such as mean, median, ranking for numerical variables and mode, ratio, and ranking for categorical variables were generated on all clustering and interpretation variables (Seth, 2021). An external consultant guided the cluster naming and interpretation process with all administrators of business units in personal finance division. Table 1 only lists simple descriptions which is summarized from the unique characteristics of the strategic and operational clusters.

Table 1. Descriptions of Strategic and Operational Clusters

Strategic (in bold format) and Operational Clusters (in bullet point format)	Unique characteristics	AUM ranking	Contribution ranking	Population size
Elite VIP	VIP customers of A-bank	1	2	4
● High-end elite VIP	Highest customer value	1-1st	1-1st	1-3rd
● Aggressive Elite VIP	Preferred high risk/high return products.	1-2nd	1-2nd	1-1st
● Conservative Elite VIP	Preferred low risk/stable return products.	1-3rd	1-3rd	1-2nd
Retired VIP	Older VIP customers	2	7	5
● Aggressive Retired VIP	Preferred high risk/high return products.	2-3rd	2-1st	1-3rd
● Conservative Retired VIP	Preferred low risk/stable return products.	2-1st	2-3rd	1-1st
● Foreign Asset Retired VIP	Preferred products settled in foreign currency	2-2nd	2-2nd	1-2nd
Potential VIP	Customers with potentials to be Elite VIP	3	4	3
● Aggressive Potential VIP	Preferred high risk/high return products.	3-1st	3-1st	3-1st
● Conservative Potential VIP	Preferred low risk/stable return products.	3-3rd	3-3rd	3-2nd
● Foreign Asset Potential VIP	Preferred products settled in foreign currency	3-2nd	3-2nd	3-3rd
Middle-Class Mortgage	Middle-class customers with mortgage and personal loan.	5	1	6
● Mortgage only middle-class	All customers with mortgage(s)	5-1st	5-1st	5-2st
● Mortgage and personal loan Middle-Class	All customers with mortgage(s) and personal loan(s)	5-3rd	5-2nd	5-1st

● Mortgage and credit card Middle-class.	All customers with mortgage(s) and credit card(s)	5-2nd	5-3rd	5-3rd
Money Demanded	Lower-income customers with mortgage and personal loan	6	3	7
● Mortgage only Money Demanded	All customers with mortgage(s)	6-1st	6-1st	6-2nd
● Mortgage and Personal loan Money Demanded	All customers mortgage(s) and personal loan(s)	6-3rd	6-2nd	6-1st
● Mortgage and credit card Money Demanded	All customers with mortgage and credit card	6-2nd	6-3rd	6-3rd
Petite bourgeoisie	Lower-middle customer looking for asset accumulation	4	5	2
● Aggressive Petite bourgeoisie	Preferred high risk/high return	4-1st	4-1st	4-1st
● Conservative Petite bourgeoisie	Preferred low risk/stable return products.	4-3rd	4-3rd	4-2nd
● Foreign Asset Petite bourgeoisie	Preferred products settled in foreign currency	4-2nd	4-2nd	4-3rd
Low Interaction	Low interaction customers with the A-bank.	7	6	1
● Foreign Asset Low Interaction	Preferred products settled in foreign currency	7-1st	7-1st	7-3rd
● Credit Card Low Interaction	Customers with credit cards only	7-2nd	7-2nd	7-2nd
● Low-Value Low Interaction	Customers with a very low frequency of interaction with A-bank	7-3rd	7-3rd	7-1st

3.2.2. The Product Affinity Prediction Model

Product affinity was defined by "*a natural liking between customers and the different products they buy*" (Baer & Chakraborty, 2013). Product affinity prediction has been widely adopted in the retail and e-commerce industries (Sarwar, Karypis, Konstan, &Riedl, 2000). The purpose of product affinity prediction is to predict customers' product preference for personalized product recommendations. A-bank's product affinity model assumed a customer's purchase decision is influenced by the customer's short-term and long-term product affinities (Guo, Cheng, Nie, Wang, Ma, & Kankanhalli, 2019). Long-term product affinity reflects a customer's product preferences rooted in the heart. The assumption of a customer's long-term product affinity can be tracked by the customer's historical transaction records and product holdings (Guo et al., 2019). Short-term product affinity reflects customer's product preferences as affected by others, such as recommendations from a wealth management advisor, from a friend or a relative, or from the media. The assumption of a customer's short-term product affinity can be tracked by a customer's recent browsing behaviors (such as on A-bank's websites) or new product transaction records (Guo et al., 2019). A customer's short-term product affinity might turn into long-term affinity with good outcomes (such as positive investment returns). Conversely, a customer's long-term product affinity might turn into short-term affinity with a bad outcome (such as a negative investment outcome).

Recency, Frequency, and Monetary (RFM) (Bult & Wansbeek, 2005)—were widely used as the target variable in product affinity prediction (Chen, Kuo, Wu, & Tang, 2009). It was used to calculate individual customers' long-term affinity scores toward the seven major personal finance products in A-bank, including mutual funds, insurance, credit cards, gold, foreign currency savings accounts, foreign currency certificates of deposit, and personal loans. Individual products' recency was defined as the period (in months) since a customer's last purchase; frequency, as the number of purchases made within the last 12 months; and monetary, as the amount of money that a customer spent within the last 12 months. Each of customers in the dataset has seven RFM scores which reflects the customer's product affinity toward seven personal finance products. The RFM scores weights of R, F, and M were based on the correlations between individual products' R, F, and M and customers' repurchase behaviors on the corresponding product within the last twelve months (Cho, Moon, Jeong, Oh, & Ryu, 2014). The RFM scores were then used as the target variable, along with other 163 input variables, to construct the predictive models. The input variables included customer's profiles, historical account activities, product holdings, and statistics of investment returns. The model predicted customers' affinity scores on the seven personal financial products. Each of the products has one predictive model which was selected via model competition. The best model was selected from—Random Forest, XG_Boosting, and Deep Neural Network—by the lowest mean squared errors in the validation dataset. That means each of customers has seven predicted RFM scores. Then the predicted scores could be sorted, to identify the top three preferred products of each individual customer.

Short-term affinity considered each customer's recent online behaviors on A-bank's websites, including both the desktop and the mobile app versions of the websites. The basic concept is like the long-term affinity model. However, the input variables were derived from customers' browsing frequencies and the target variables were products' RFM scores of the recent six months. The short-term product affinity might have squeezed out the top three product affinity if any predicted scores of short-term product affinity was higher than the top three long-term product affinities.

3.3. Management Process

To advocate the necessary change management in banks to successfully leverage big data analytics, we propose the following framework and discuss strategies for changes.

3.3.1. The Framework of the Automatic Omni-Channel Implementation

Figure 3 shows the framework of the automatic omni-channel implementation by combining the results of operational clustering and product affinity. First, all customers were labeled with the strategic cluster, the operational cluster, and the top three preferred products every month via autoscoring. These labels were synced with the following systems: (1) all websites; (2) the branch's front desk system; (3) the bank's call center; and (4) the wealth management system for VIP customers. Possible methods of interaction included inbound (customer self-initiated) or outbound (bank-initiated). Non-VIP customers mainly used the low-cost channels, such as websites and the branch's front desk system, for interactions. For VIP customers, the major interaction channel became the bank's wealth management advisors. A customer might receive personalized product recommendations while surfing on one of A-bank's websites based on "clustering + product affinity" approach. When a customer reaches any A-bank branch's front desk, the banker can offer personalized product recommendation as he/she retrieves customer's account information from the front desk system. The banks can recommend products while dealing with the customer's request. The same top three preferred product recommendations are stored in both the call-center and the wealth management systems. Either the call-center agent or the wealth management advisor can offer personalized recommendations when a customer calls in or walks in, or they can call out to customers who have the highest probability of buying specific financial products.

This is an author-produced, peer-reviewed version of this article. The final, definitive version of this document can be found online at *Journal of Enterprise Information Management*, published by Emerald Publishing Ltd. Copyright restrictions may apply. https://doi.org/10.1108/JEIM-05-2020-0176.

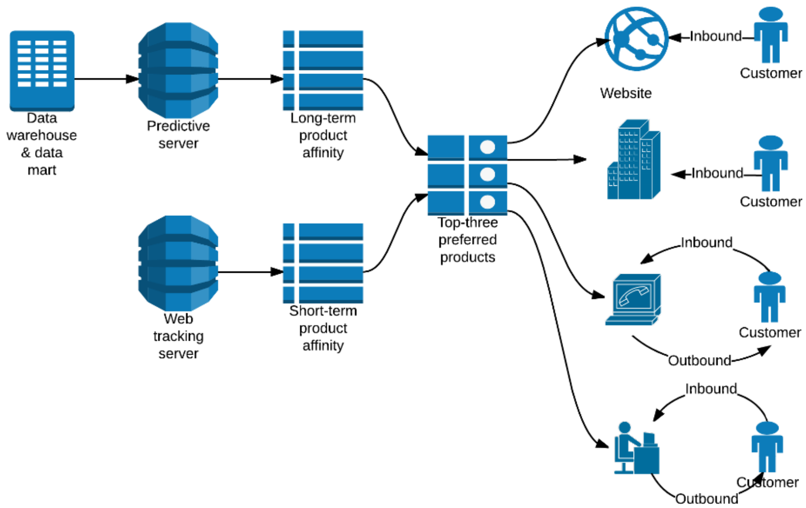

Figure 3. The auto omni-channel marketing framework

3.3.2. The Outbound Campaigns via Wealth Management Advisors

The practical model validations were implemented via the omni-channel marketing. Since some of the channels were only recently implemented, the authors chose to report the outbound campaigns conducted by the wealth management advisors. Table 4 shows the target customers, the model adopted for customer selection, the implementation duration, the number of branches involved, and the average response rate. The target financial products of these specific campaigns were mutual funds and insurance. The goal was to prove the effectiveness of personalized recommendation (to introduce the right products to the right person). Therefore, these campaigns did not provide any special offers to customers. In addition, the product affinity model scores customers based on their likelihood to purchase, so that marketing campaigns can be targeted towards those customers who will maximize the response rate. However, the division of personal finance wanted to prove that the response rate was triggered by the models, not by the customers who would respond positively with or without campaign activities. Therefore, in addition to the clustering and product affinity predictions, the target customers had to be zero mutual fund and insurance product holding or not having purchased these two products in the past six months. The total sales of the three waves were about 0.3 billon US dollars. These campaigns are briefly described below.

The first wave aimed to validate the effects of two-stage clustering. Therefore, the first wave only involved one third of the branches. These branches were selected because these branch managers volunteered to join the first-wave implementation. In the first round, the analytics team tested whether the wealth management advisors could activate conservative high AUM customers. Therefore, customers named Conservative Elite VIP, Conservative Retired VIP, and Foreign Asset Potential VIP were selected as the target customers. The historical campaign response rate was about 2.4% (baseline), and the response rate of the first wave was 3.1%. The leader of the Personal Finance Division was very happy with the results because: (1) the target customers were conservative customers, but the response rate was higher than the baseline, and (2) The campaign saved lots of time and effort for the wealth management advisors. However, the analytics team also found that the conservative customers accepted only low-risk products, like insurance.

Because of the bank's successful experience during the first wave, all the branches were involved in the second wave, and both the clustering and the product affinity models were adopted in the customer selection. The results of the first wave indicated that conservative customers would accept insurance products only. Therefore, the selection of customers focused on the following operational clusters: Aggressive Elite VIP, Conservative Elite VIP, and Aggressive Retired VIP. The campaign was very successful, with a 4.9% response rate.

The third wave added two more operational clusters: Aggressive Potential VIP and Foreign Asset Potential VIP. The response rate increased to 6.5%.

Table 2. An overview of Practical Campaigns Using the Models

Champaign	Target customers	Model adopted	Duration	Branches	Campaign response rate (baseline is about 2.4%)
1st Wave	All the customers were zero product holding or had not purchased any mutual fund and/or insurance in the last six months.	Model: Two-stage clustering Clusters: Conservative Elite VIP, Conservative Retired VIP, and Foreign Asset Potential VIP	Four months	1/3 branches	3.1%
2nd Wave	Same as the first wave	Models: Two-stage clustering and Short-term & long-term product affinity Clusters: Aggressive Elite VIP, Conservative Elite VIP, and Aggressive Retired VIP	Three months	All	4.9%
3rd Wave	Same as the first wave	Models: Two-stage clustering and Short-term & long-term product affinity Clusters: Aggressive Elite VIP, Conservative Elite VIP, Aggressive Retired VIP, Aggressive Potential VIP, and Foreign Asset Potential VIP	Three months	All	6.5%

Table 3 further compares the response rates and the rankings of total contributions by operational clusters in the first three waves of implementations. Because the first wave utilized the clustering analysis only and selected two conservative clusters as the target, the outcome was only 29% higher than the baseline. The rest of the two waves had 104% and 171% higher response rates, respectively, than the baseline, when incorporating aggressive customers. The results also show that the wealth management advisors and the personal finance division improved their practice via these campaigns, since response rates were higher than the previous wave within the same cluster. The results also show that high AUM customers (Elite VIP and retired VIP) made the highest total contributions, although the response rates might not be the highest.

Table 3. Comparing the response rates and rankings of total contributions

Operational cluster	First wave's response rate (rankings of total contributions)	Second wave's response rate (rankings of total contributions)	Third wave's response rate (rankings of total contributions)
Aggressive Elite VIP		11.4% (1st)	12.0% (2nd)
Conservative Elite VIP	2.3% (1st)	2.5% (2nd)	3.4% (1st)
Aggressive Retired VIP		7.2% (3rd)	10.4% (3rd)
Conservative Retired VIP	1.7% (2nd)		
Aggressive Potential VIP		6.1% (4th)	8.9% (5th)
Foreign Asset Potential VIP	5.4% (3rd)		6.3% (4th)
Overall response rate (baseline: 2.4%)	3.1% (29% higher than the baseline)	4.9% (104% higher than the baseline)	6.5% (171% higher than the baseline)

Figure 4 shows the validation of the product affinity model with campaign results. The solid line illustrates the actual mutual fund response rate versus the predicted mutual fund affinity probability, and the dash line shows the actual insurance response rate versus the predicted insurance affinity probability. Basically, the results indicate that the product affinity model is fairly accurate, since the actual response rates align with the predicted probabilities. Because the response rates of insurance are significantly smaller (i.e., a smaller sample size), the dash line shows a similar trend to that of the mutual fund, but with some irregular peaks. Figure 4 also shows that the response rates of mutual fund are significantly higher than insurance, by four to five times.

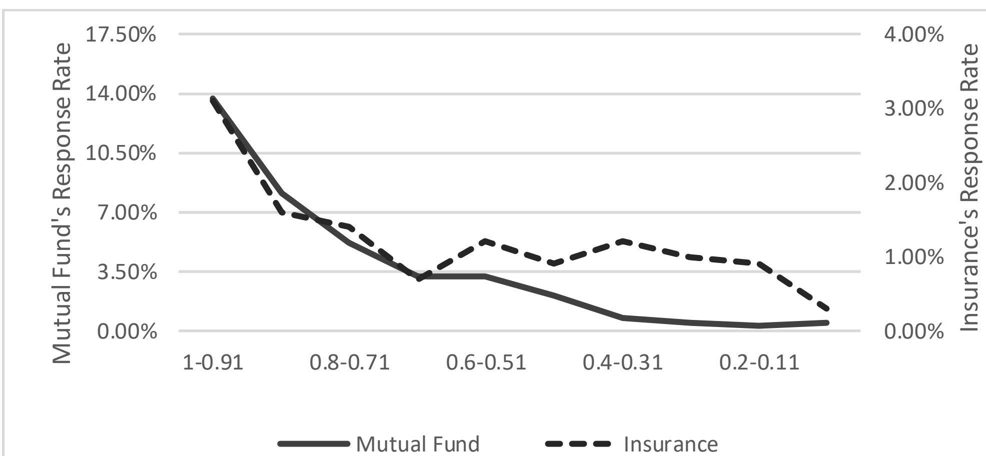

Figure 4. Response rate by product affinity probability

The three waves of the campaign lasted for 10 months. The migration (by comparing a customer's cluster label before and after campaign implementations) of strategic clusters could serve as evidence for examining and adjusting market strategies. Table 4 only lists the migrations of four major strategic clusters in the campaign implementations. Each of the clusters had an increase of customers from other clusters (migrating in) and had a loss to other clusters (migrating out). Here is the summary of the cluster migrations after the campaigns: (1) Elite VIP: The average AUM was even, but the average of contribution increased 1.4%. The major migrate-in and migrate-out cluster was Potential VIP. Because the migration in was larger than the migration out, the total number of Elite VIP increased 3.9%. (2) Retired VIP: The average AUM decreased 0.3% and the average contributions decreased by 0.6%. The largest migrate-in and migrate-out clusters were Potential VIP and Elite VIP. Because the net migration was positive, the total number of Retired VIP increased 1.1%. (3) Potential VIP: The average AUM and contributed increased 0.7% and 3.7%, respectively. The major migrate-in and migrate-out clusters were Petite bourgeoisie and Elite VIP. However, due to the net migration being negative, the total number of Potential VIP decreased 2%. (4) Petite bourgeoisie: The average AUM and contribution increased by 7.2% and 4.1%, respectively. The largest migrate-in customers were new customers. The largest migrate-out cluster was Potential VIP. Due to the net migration being negative, the total number of customers decreased 7.7%.

Table 4. Migration of Strategic Clustering for Campaigns

Strategic clusters	Average AUM	Average Contribution	Migrate out	Migrate in	Number of Customers
Elite VIP	Even	+1.4%	Potential VIP	Potential VIP	+3.9%
Retired VIP	-0.3%	-0.6%	Elite VIP	Potential VIP	+1.1%
Potential VIP	+0.7%	+3.7%	Elite VIP	Petite bourgeoisie	-2%
Petite bourgeoisie	+7.2%	+4.1%	Potential VIP	New customers	-7.7%

3.3.3. Interview with Wealth Management Advisors

Five branches were selected in the first and the third wave implementations for interview. These branches were selected because their campaign performance was the highest or the lowest in first or the third waves. Four major questions guided the process of interview. The analytic team found wealth advisors interpreted the analytics results in their own ways. Low performing advisors interpreted the results from different perspectives and their implementation strategies were conflicted with the original implement plan. Table 5 below summarizes responses from high and low performing wealth management advisors. Some negative effects were found after interviewing the wealth management advisors. Most issues identified during the implementation were related to incongruences among different stakeholder groups and resistance toward change. Several strategies were adopted to alleviate the negative effects.

Table 5. Interview Results

Q1: How did you start the implementation when you received the list of potential customers?	High performers: • Skimmed through the list of potential customers. Constructed a plan to deal with different scenarios when talked with the target customers (the third wave). • Started with the beginning of the list (the first & the third waves). • Called at least three times if the customer did not answer the call (the first & the third waves). Low performers: • Skim through the list of potential customers. Construct a plan to deal with different scenarios when talked with the target customers (the third wave). • Selected customers based their own experiences (the first wave). • Call at the least three times if the customer did not answer the call (the third waves).
Q2: Were there any difficulties encountered during the process of implementation?	High performers: • Customers refused to take any solicitation calls. However, the situation is getting better after the first two waves (the third wave). Low performers: • Without any rewards or promotions associated with the campaign, it is hard to start the conversations with the target customer. Especially, when the customer is new to the wealth management advisor (the first wave). • The prediction is not accurate because the target customers with higher predicted probabilities did not respond the campaign (the first wave).
Q3: Are there any additional supports needed to improve your performance next time?	High performers: • No, the analytics results were comprehensive and useful (the third wave). Low performers: • Please limit the list to high-asset customers next time (the first wave).
Q4: Are there any additional comments?	High performers: • The approach assisted me to identify customers with high potentials. Saving me lots of time and efforts. I look forward to receiving more new customers to enlarge my customer basis (the third wave). Low performers: • Is there any system which allows me to create a list of potential customers on my own? (the first wave)

4 Discussion

4.1. Generalizable Explanations with TCT

With rapid digital transformation, it becomes unclear about the role of TCT in understanding when it is more efficient for a transaction between two parties to occur within the market or within an organization. In addition, it is unclear whether TCT can still explain the competitions between traditional companies and their online competitors.

Traditional banks face many challenges from companies including FinTech companies and online mutual fund platforms. There are currently three online mutual fund platforms in Taiwan. Nagle, Seamans, & Tadelis (2020) summarized three unique characteristics of digitally-mediated transactions: (1) the ability for a digitally mediated transaction to be free of charge, (2) the amount of private or personally identifiable data about parties that is conveyed either before or during the transaction, and (3) reputation mechanisms that are used to overcome information asymmetry. Online companies have become key players in many industries, such as media and retail industries. These online companies can provide cheaper or even free products and services than traditional companies and thus can acquire the majority of the market shares. However, online mutual fund platforms have not showed up as strong competitors for traditional banking industries in Taiwan. Comparing with other industries, the unique contexts in banking industry provide competitive advantages to lessen challenges from online mutual fund platforms.

The major "marketing hooks" for online mutual fund platforms are cheaper transaction fees and more choices of mutual fund products. However, Securities and Futures Institute (2018) found that 89.4% of respondents in Taiwan relied on wealth management advisors to suggest mutual fund products and when to buy and sell mutual funds. 91.7% of respondents selected the mutual fund performance as the most important factor when making investment related decisions. Only 46.9% of respondents felt that transaction fee is important. Most people hope to obtain professional consultation before making purchase decision. That means the major marketing hooks are becoming less attractive to customers.

The online mutual fund platforms reply on customer browsing behaviors to construct similar customer clustering model and product affinity models, similar to what traditional commercial banks do. However, Taiwan's Financial Supervisory Committee has strict regulations to protect personal private information. Therefore, the online mutual fund platforms are limited to the use of personal private information for selling advertising or other services. Comparing with online mutual fund platforms, traditional banks track and store all customers' information in personal finance. Traditional banks can still identify high value customers from their other product transactions or holdings, even these customers never had any mutual fund transactions.

Finally, unlike e-commerce companies, the reputation mechanism on the mutual fund platforms is not intuitive to customers. The most common reputation mechanism is a rating system with various indictors to evaluate performance of individual mutual fund products. Customers may be overwhelmed by numerous jargons and abbreviations. Although many mutual fund platforms provide free robot advisor services to monitor individual customers' investment portfolios, only 59.3% of Taiwanese are interested in using online tools for wealth management (CITI bank, 2018). The percentage increased to 77.8% (on average; 80.7% for high-end customers) if they could discuss with wealth advisors before making their investment decisions.

Based on the TCT, if a customer has enough financial knowledge, they would select the online mutual fund platform for lower transaction cost. However, as most customers do not have sufficient financial knowledge or time to select mutual fund products or monitor personal portfolios, outsourcing to wealthy management advisor is a common approach. Outsourcing is associated with different kinds of pre- and post-transaction costs, such as searching, contracting, monitoring, and adaption. When A-bank's wealth management advisors contacted potential customers based on the models' results, they also saved these customers' searching cost at the same time. Due to the unique characteristics of financial products, the market share of personal wealth management is currently not dominated by FinTech companies.

4.2. The Adoption of Big Data Analytics

This case study enriches our understanding of the approach and factors that affect the implementation of big data analytics in banks. The proposed framework in figure 1 allows organizations to develop roadmaps for implementing big data analytics in their organizations and address concerns from stakeholders including various groups and practitioners. It highlights the importance of non-technical factors in influencing the implementation success of big data analytics in organizations.

The results of this exploratory case shed light on the implementation and adoption level of big data technologies in organizations. The results from the multiple-wave implementations can be used to seed theory development for the dynamics of big data adoption, to identify the factors that affect the dynamics of big data implementation. As different stakeholders in the organizations engage in changing routines and processes enabled by big data analytics technology, inconsistencies and incongruences in stakeholders' technological frames will be playing a role in the change process. Future studies could investigate multiple-wave implementations of the adoption of big data technologies on the impact of different stakeholders in different stages or settings.

The TCT can further explain implementation results of different stakeholders. From the perspective of A-bank, considering all the cost associated with data collection, private data protection, model construction, and model maintenance, the transaction (i.e. big data adoption) should occur within the organization, instead of outsourcing to an external big data service companies. However, since this is all new for A-bank, to lower the level of uncertainty, A-bank hired an external consultant to construct these models then internal IT and the analytic team are in charge of data collection, data protection, and model maintenance. The format of partial outsourcing is very common in the banking industry.

From the perspective of wealth management advisor, the campaigns replied on analytic results of customer clustering and product affinity prediction. These two models utilized data collected from the data warehouse to provide a holistic overview of the customers. The analytic results lowered wealth management advisor's bounded rationality and task complexity. Customer clustering provides a convenient way for customer selection. The selected clusters for campaigns were customers who were regarded to generate higher short-term or long-term revenues. Product affinity prediction further decreased costs of uncertainty and task complexity. All wealth management advisors can develop personalized strategies based on the product affinity prediction to manage individual customers. After multiple-wave implementations, the information processing cost was further decreased when wealth management advisors got familiar with the whole process. The effect can be observed via the rising response rates across multiple waves.

From the perspective of customers, working with personalized wealth management advisor can lower customer's search costs and time (Mooi & Ghosh, 2010). Because wealth management advisors presented each of campaign customers personalized recommendations based on product affinity prediction, the mechanism lowered time costs on both parties and boosted the chance of efficient transaction. However, the response rates in Table 3 show the importance of long-term relationship management. Millman and Wilson (1996) defined key customers as customers who fit the 80/20 rule. In the case study, the key customers are Elite VIP customers. Because these customers have recognized A-Bank's CRM effort and benefits, wealth management advisors are easier to achieve efficient transactions with Elite VIP. From aspect of customer acquisition cost, the order from high to low is Petite bourgeoisie > Retired VIP > Potential VIP > Elite VIP. The migrations of strategic clusters after multiple-wave campaign shows customer's perceived benefits might increase after more interactions occurred between customers and wealth management advisors. For example, Potential VIP might migrate to the Elite VIP, in terms of customer loyalty and revenue growth. However, if the customer's perceived benefits were lower than the customer's expectation, then the customer might migrate out.

4.3. Practical Implications

The above framework explains the roadmaps of the big data adoption and why customers accepted the personalized precision marketing. However, during the process of implementation, not all the results occurred as expected. First, the interview results showed that different people interpreted the analytic results differently and had different expectations. In the first wave, some wealth management advisors only contacted customers who were selected by personal experiences. The additional filtering mechanism increased difficulties on tracking the actual effects of the predictive models. Therefore, one major challenge is how to help decision makers better understand and interpret the analytic results. Several strategies had been adopted to alleviate the negative impact. For example, the analytic team provided training to explain the models' concepts and potential applications. The goal was to correct stakeholders' perceptions and increase their confidence on big data analytics.

Second, some experienced wealth management advisors and administrators had doubt about the power of big data analytics. These advisors felt they can achieve the assigned KPI without data analytics. They also attributed the campaigns' success to personal efforts instead of big data analytics. To alleviate negative effects, the analytic team provided stakeholders analytic results to demonstrate models' effectiveness. In addition, a ceremony was conducted to recognize high performing wealth management advisors. These advisors were also asked to share their success strategies in the ceremony. According to TFR, IT-based projects may also exhibit frame incongruences between key stakeholder groups at various stages. It is likely to need long-term efforts to reconcile incongruences and help these advisors trust and adopt data analytics.

5. Recommendations for Big Data Analytics Implementation

5.1. The Generation of Meaningful Analysis Results

In most companies, the big data analytics team serves as support staff to provide analysis results for better management and decision making. The administrators and the marketing sales force can be the major hurdles of resistance in the process of reform, since they may not understand or trust the underlying techniques well and heavily rely on personal prior success experiences to make judgment or decisions (Crawford, 2013). It is essential to make the analysis of results meaningful to stakeholders. For example, in clustering analysis, the big data analytics team selected key variables for strategic clustering and presented the clustering analysis results with the dimensions of contribution and the AUM. It might not be the best modeling approach, but the results were very meaningful to the staff and other stakeholders involved in big data analytics-related discussions. Since there is no universal standard to evaluate practical value of clustering analysis, meaningful and actionable insights can be one of the most important goals. This is a practical contribution to the measurement of big data implementation outcomes in the banking industry.

Another important performance indicator is the response rate of the campaign implementation (Mirbagheri & Hejazinia, 2010). In this case study, the selection of customers was limited to customers who were zero product holding or had not purchased any mutual funds and/or insurance during the prior six months. Therefore, the incremental sales can be regarded as the models' effectiveness. However, in other big data analytics projects, selecting an appropriate baseline to highlight the effectiveness of the model can sometimes be a tougher task than the analysis itself.

5.2. From Small-Scale to Large-Scale Implementation

Several studies have recommended smaller scale implementation to examine every internal and external key points in the implementation process (Sivarajah et al., 2017). The first wave of campaign implementation involved one third of the voluntary branches. Those branch managers can be considered as the early adopters. Because those branch managers volunteered to participate in the test run, the analytics team was sure that they would implement the campaign with their full efforts. Not only would these managers share their personal experiences with other branch managers, but the positive outcomes would serve as the best evidence to increase the confidence of other branch managers who has not participated in the first round. The multiple-wave implementations allowed all involved administrators to get familiar with related processes and to observe customers migrations among clusters (see Tables 3 and 4).

5.3. The Selection of Marketing Channels

Marketing cost is another important consideration in campaign implementation. Table 3 shows that the high response rate does not equal to high total contributions. That is why omni-channel is a cost-effective approach in campaign marketing. Low-cost channels, like websites and mobile app push ads, are suitable for low-profit margin products or customers (Li & Du, 2017), and wealth management advisors are reserved for high-profit generation customers (Lau, Chow, & Liu, 2004), although these customers might have significantly lower response rates. Moreover, the analytics team found that customers who visited corresponding web pages after an advisor's introduction had the highest response rate, across all the campaign clusters. Finally, long-term product affinity is rooted deeply in the past. When historical behaviors classified customers as conservative, the campaign results show that these customers accepted conservative products only (see Table 3).

6. Conclusions

Evidence-based practices are needed to guide banks and other types of organizations to implement big data analytics for enhancing their decision making and increasing their competitiveness. This case study demonstrates how the proposed framework in figure 1 can be used to guide the implementation of big data analytics projects and help organizations successfully address challenges in the context of leveraging big data analytics.

There are several limitations with this study. First, we only provided one in-depth case study for our research. More case studies or a large-scale survey study from multiple data sources should be done in the future to validate the findings. Secondly, our exploratory case study only reported cluster analysis and product affinity models. Future research should examine other big data analytics methods and make a comparative study to advance the theories and practices. Thirdly, this study reported some non-technical and technical factors for successful implementation of big data analytics projects in banks. More empirical research on critical success factors is recommended. In addition, future research can focus on the best practice of big data analytics for increasing customer values and leveraging marketing campaign.

References

Alles, M. G. (2015). Drivers of the use and facilitators and obstacles of the evolution of Big Data by the audit profession. *Accounting Horizons*, *29*(2), 439-449.

Alghofaili, Y. (2020). Interpretable K-Means: clusters feature importances. Towards Data Science. Retrieved from https://towardsdatascience.com/interpretable-k-means-clusters-feature-importances-7e516eeb8d3c

Aubert, B., & Weber, R. (2001). *Transaction Cost Theory, The Resource Based View, and Information Technology Outsourcing Decision*. A Re-Examination of Lacity Et Al's Findings, HEC Montreal Yayınları, Kanada.

Baer, D., & Chakraborty, G. (2013). Product affinity segmentation using the doughnut clustering approach. *In Proceedings of the SAS Global Forum 2013 Conference*, Cary, NC.

Baker, J. (2012). The technology–organization–environment framework. In *Information systems theory* (pp. 231-245). Springer, New York, NY.

Bult, J. R. & Wansbeek, T. J. (2005). Optimal selection for direct mail. *Marketing Science*, *14*(4), 378–394.

CITI Bank (2018, June 20). Customers prefer combining online investment tool with professional consultation. MoneyDJ News. Retrieved from https://www.moneydj.com/kmdj/news/newsviewer.aspx?a=2b08a897-df6c-4c69-a546-9a82c323d7f0

Chedrawi, C., Atallah, Y., & Osta, S. (2020). Big Data in the Banking Sector from a Transactional Cost Theory (TCT) Perspective—The Case of Top Lebanese Banks. In *ICT for an Inclusive World* (pp. 391-405). Springer, Cham.

Chen, Y. L., Kuo, M. H., Wu, S. Y., & Tang, K. (2009). Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data. *Electronic Commerce Research and Applications*, *8*(5), 241-251.

Cho, Y. S., Moon, S. C., Jeong, S. P., Oh, I. B., & Ryu, K. H. (2014). Clustering method using weighted preference based on RFM score for personalized recommendation system in u-commerce. In *Ubiquitous information technologies and applications* (pp. 131-140). Springer, Berlin, Heidelberg.

Crawford, K. (2013). The hidden biases in big data. Harvard Business Review. Retrieved from https://hbr.org/2013/04/the-hidden-biases-in-big-data

Collica, R. S. (2017). Customer segmentation and clustering using SAS Enterprise Miner. SAS Institute.

Davis, C.J., & Hufnagel, E.M. (2007). Through the Eyes of Experts: A Sociocognitive Perspective on the Automation of Fingerprint Work. *MIS Quarterly*, 31 (4), pp. 681–703.

Davidson, E. (2006). A Technological Frames Perspective on Information Technology and Organizational Change. *The Journal of Applied Behavioral Science*, 42(1), pp. 23–39.

De Laat, P. B. (2018). Algorithmic decision-making based on machine learning from Big Data: Can transparency restore accountability? *Philosophy & Technology*, *31*(4), 525-541.

Dudek A. (2020) Silhouette Index as Clustering Evaluation Tool. In: Jajuga K., Batóg J., Walesiak M. (eds) *Classification and Data Analysis*. SKAD 2019. Studies in Classification, Data Analysis, and Knowledge Organization. Springer, Cham. https://doi.org/10.1007/978-3-030-52348-0_2

Dolci, P. C., Maçada, A. C. G., & Grant, G. G. (2014). IT investment management and information technology portfolio management (ITPM): Brazilian case studies. *Journal of Enterprise Information Management*, Vol. 27 No. 6, pp. 802-816.

Elzamly, A., Hussin, B., Naser, S. A., Khanfar, K., Doheir, M., Selamat, A., &Rashed, A. (2016). A New Conceptual Framework Modelling for Cloud Computing Risk Management in Banking Organizations. *International Journal of Grid and Distributed Computing*, 9(9), 137-154.

Esmin, A. A., Coelho, R. A., &Matwin, S. (2015). A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data. *Artificial Intelligence Review*, *44*(1), 23-45.

Guo, Y., Cheng, Z., Nie, L., Wang, Y., Ma, J., & Kankanhalli, M. (2019). Attentive long short-term preference modeling for personalized product search. *ACM Transactions on Information Systems (TOIS)*, *37*(2), 1-27.

Haenlein, M., Kaplan, A. M., &Beeser, A. J. (2007). A model to determine customer lifetime value in a retail banking context. European Management Journal, 25(3), 221-234.

Heide, J. B. (1994). Interorganizational governance in marketing channels. *Journal of marketing*, *58*(1), 71-85.

Hendalianpour, A., Razmi, J., &Gheitasi, M. (2017). Comparing clustering models in bank customers: Based on Fuzzy relational clustering approach. *Accounting*, *3*(2), 81-94.

Jones, S. (2012). Understanding t-government progress and issues: a case study of a United Kingdom local authority. *Journal of Enterprise Information Management*, Vol. 25 No. 6, pp. 605-618.

Keogh E. & Mueen A. (2017) Curse of Dimensionality. In: Sammut C., Webb G.I. (eds) *Encyclopedia of Machine Learning and Data Mining*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4899-7687-1_192

Khandare, A., & Alvi, A. S. (2016). Survey of improved k-means clustering algorithms: improvements, shortcomings and scope for further enhancement and scalability. In Information Systems Design and Intelligent Applications (pp. 495-503). Springer, New Delhi.

Lau, K. N., Chow, H., & Liu, C. (2004). A database approach to cross selling in the banking industry: Practices, strategies and challenges. *Journal of Database Marketing & Customer Strategy Management*, *11*(3), 216-234.

Klimis, G. M. (2007). *Electronic Markets*. The Blackwell Encyclopedia of Sociology.

Li, F., & Du, T. C. (2017). The effectiveness of word of mouth in offline and online social networks. Expert Systems with Applications, 88, 338-351.

Ma, J., Baer, D., & Chakraborty, G. (2015). *Statistical evaluation of the doughnut clustering method for product affinity segmentation*. Retrieved from https://business.okstate.edu/site-files/docs/analytics/3150-2015.pdf

Maiste, P. (2018). Improving marketing programs with predictive analytics. *LITYX*. Retrieved from https://lityx.com/improving-marketing-programs-predictive-analytics/

Marshall, A., Tang, L., & Milne, A. (2010). Variable reduction, sample selection bias and bank retail credit scoring. *Journal of Empirical Finance*, *17*(3), 501-512.

Melchiotti, R., Gracio, F., Kordasti, S., Todd, A. K., & de Rinaldis, E. (2017). Cluster stability in the analysis of mass cytometry data. *Cytometry Part A*, *91*(1), 73-84.

Michael, M., & Lupton, D. (2016). Toward a manifesto for the 'public understanding of big data.' Public Understanding of Science, 25(1), 104–116.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, 98, 261–276.

Millman, T., & Wilson, K. (1996). Developing key account management competences. *Journal of Marketing Practice: Applied Marketing Science*, *2*(2), 7-22.

Mirbagheri, S., &Hejazinia, M. (2010). Mobile marketing communication: learning from 45 popular cases for campaign designing. *International Journal of Mobile Marketing*, *5*(1), 175-192.

Mooi, E. A., & Ghosh, M. (2010). Contract specificity and its performance implications. *Journal of Marketing*, *74*, 105–120.

Nagle, F., Seamans, R., & Tadelis, S. (2020). Transaction cost economics in the digital economy: A research agenda. *Harvard Business School Strategy Unit Working Paper*, (21-009).

Naveira, C. F., Jacob,I., Rifai, K., Simon, P. & Windhagen, E. (2018). Smarter analytics for banks. Retrieved on April 23, 2020 at https://www.mckinsey.com/industries/financial-services/our-insights/smarter-analytics-for-banks

Ngai, E. W., Xiu, L., & Chau, D. C. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, *36*(2), 2592-2602.

Orlikowski, W.J., and Gash, D.C. (1994). Technological Frames: Making Sense of Information Technology in Organizations. *ACM Transactions on Information Systems*, 12(2), pp. 174–207.

Ragin, C. (1992). Casing and the Process of Social Inquiry. In Ragin, Charles C. and Becker, Howard S. (eds), *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press, pp. 217-226.

Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, 38(1), 187-195.

Sambandam, R. (2003). Cluster Analysis Gets Complicated Collinearity is a natural result of clustering. So how can researchers get around it?. *Marketing Research*, *15*(1), 16-21.

Sanader, D., & Marko, L. (2017). Marketing analytics in the function of decision making in banks. *Časopis za ekonomiju*, 107.

Sarwar, B., Karypis, G., Konstan, J., &Riedl, J. (2000). Analysis of recommendation algorithms for e-commerce. In Proceedings of the 2nd ACM conference on Electronic commerce (pp. 158-167). ACM.

Seth, N. (2021). Profiling Market Segments using K-Means Clustering. *Data Science Blogathon*. Retrieved from https://www.analyticsvidhya.com/blog/2021/01/profiling-market-segments-using-k-means-clustering/

Shirazi, F., &Mohammadi, M. (2019). A big data analytics model for customer churn prediction in the retiree segment. *International Journal of Information Management*, 48, 238-253.

Sivarajah, U., Kamal, M. M., Irani, Z., &Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, *70*, 263-286.

Srinivasan, S., & Kamalakannan, T. (2018). Multi criteria decision making in financial risk management with a multi-objective genetic algorithm. *Computational Economics*, *52*(2), 443-457.

Watjatrakul, B. (2005). Determinants of IS sourcing decisions: A comparative study of transaction cost theory versus the resource-based view. *The Journal of Strategic Information Systems*, *14*(4), 389-415.

Williamson, O. E. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications, A Study in the Economics of Internal Organization*. New York: Free Press.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press.

Yeoh, W., & Koronios, A. (2010). Critical success factors for business intelligence systems. *Journal of computer information systems*, 50(3), 23-32.