Digital Mobile Services for Consumer Banking
Conceptualization and Implementation of Location-based Consumer Credits

Wenzel-Ruelberg, Clemens; Baiyere, Abayomi; Lohmann, Patrick

Digital Mobile Services for Consumer Banking:
Conceptualization and Implementation of Location-based Consumer Credits

Clemens Wenzel-Ruelberg
Copenhagen Business School
clemens@ruelberg.com

Abayomi Baiyere
Copenhagen Business School
aba.digi@cbs.dk

Patrick Lohmann
zeb
Patrick.lohmann@zeb.de

Abstract

Interactions of banks with their customers are increasingly shifting to web and mobile channels. Being at risk of losing the role of the customer agent to FinTechs and digital challenger banks, incumbent banks are seeking ways to exploit technologies such as mobile phones as channels to generate insights and to sell products that their customers need. For many banks the potential of mobile banking to individualize products and personalize services from information collected by the mobile device' sensory components are largely untapped. In this paper, we draw on design science research in exploiting spatio-temporal information to build an algorithmic model to target customers with credit offers. While past research aimed to solve similar problems mainly through customer segmentation, our approach demonstrates the benefits of having a transparent and interpretable decision model for each individual customer. Our artifact enables the development of digital products and services, without large-scale, often unavailable data.

1. Introduction

Remaining low interest rates, ongoing technological advances, and increased competition are forcing banks to rethink their business and operating models to stay profitable. The ongoing integration of digital technology into banking is omnipresent. The convergence of previously disparate developments such as connected mobile devices, wireless networks, web technologies, and the internet of things creates opportunities for an unparalleled customer orientation, entrepreneurial spirit, and digital innovation [2].

To surf on these trends and react to the shifting customer expectations, banks are looking for ways to participate more deeply in the life of their customers, providing purposeful services through the increasingly important mobile channel. Banking is a relationship business and to remain a close relationship with their customers, it is vital for them to learn about and proactively reach out to their customers. At the same time, it is critical that customers remain receptive to such interactions, requiring a balance between an overflow of encounters and no interaction at all.

Any bank has access to its customer information such as incoming and outgoing transactions or other information that are implicit and hidden in transactional data. This lays the groundwork for analytical models that use statistical methods and machine learning to computationally create profiles and obtain valuable information about customer wishes and wants, and to predict customer behavior upon that. One application area to exploit the value of data is in the provisioning of financial products such as consumer credits.

Our research objective is to explore how banks may leverage the affordance of location based technologies in unlocking new business opportunities. Hence, the specific purpose of this paper is to develop an algorithmic model that predicts when a client likely needs a consumer credit based on his/her location. While the model is intended to use conventional information such as transaction data from customers' bank accounts and information about their preferences, another type of data is added – location data. Depending on a customer's geographical position and their preferences, targeted credit offers can be sent directly to the customer's smartphone. In a situation in which the customer could make use of consumer credit, an individual offer is sent. The client can then either accept or reject it. For the development of this prototype, a German bank provided us with anonymized customer data. Prior to this study, the customers have consented to share their data for research and testing purposes. A later implementation of our model must succeed with the consent of a customer given to all participating parties prior to signing up. This is a necessary step in order to use the model in accordance to the GDPR in the EU. While our model is developed for credit offerings, it can easily be adapted to other financial services contexts where the client may be advised.

To clarify the practical issue that our DSR study addresses, we outline three main benefits of the pro-

posed solution to the key stakeholders. First, our artifact is intended to equip banks with the confidence to initiate a digital interaction by offering more targeted products that the customer is likely to accept. Compared to other forms of consumer credits (e.g. credit cards), it is the bank and not the customer who initiates the lending process by sending an offer in a specific situation. If successfully implemented, this can ultimately result in increased revenue for the bank, as more credits are emitted. Second, many customers perceive financial products as complex and opaque, and hence they may not always know which products exist on the market and which specific one may best solve their problem. Having a more sophisticated understanding of the behavior of their clients can help banks provide more personal advice and is therefore of great value to both the customer and the bank. Third, customers, on the other hand, benefit from having another payment option next to other installment plan payment methods such as credit cards or in-store financing options. In-store financing options usually require lengthy sign-up procedures and background checks at each store individually. When paying with credit card, though, a credit card company is sitting between the customer and the bank with often higher interest rates. Our proposed model does not intend to involve third parties in this concern.

Adding a spatio-temporal dimension to conventional profiling models, we argue that the artifact is more accurate in understanding and predicting customer desires and needs. It offers a starting point for a more controlled and incremental ingestion of digital technologies and artificial intelligence into the modus operandi of banks which are traditionally more risk averse than FinTechs and challenger banks when it comes to adapting the business and operating model because of their legacy.

Through five expert interviews, we identified three objectives for our artifact, which serve as our evaluation measures:

a) Only send customer-specific (targeted) offers
b) Avoid an overflow of offers for a specific customer (no spamming)
c) Evaluate the customer responses and consider the feedback for decisions (continual learning)

First, only targeted offers shall be sent to achieve higher acceptance rates. The purpose of a credit offer needs to meet the preferences of the individual customer, e.g. that the credit is intended for certain non-financial products in stores. Specific historic buying activities need to be considered. Even if tailor-made offers are sent, the risk of overflowing customers with offers still exists. The model therefore must adapt the frequency of offers. Location data shall be used, as it provides spatial and temporal information of a customer. In combination with buying behavior in specific situations, opportunities are identified to adjust offer frequencies. When a credit offer is sent, the customer can accept or reject it. This feedback information in turn provides valuable insights for further profiling and must be included in the model to learn from customer feedback and to continuously improve.

Applying design science research, we proceed with a literature review about relevant themes and existing models for customer profiling within consumer banking. We then describe our approach to design, build, and test our model. We evaluate if the artifact meets the objectives and hence its practical utility. We discuss the results, reflect on our method and compare the artifact with existing solutions.

2. Literature Review

2.1. Mobile Computing and Big Data

Location-based services (LBS) were mentioned for the first time by the US government in a 1996 mandate that required mobile-network operators to locate emergency calls precisely enough to inform first responders. The positioning algorithms used to determine a person's location were costly and did not receive much attention. The introduction of GPS-capable mobile devices 3G broadband wireless service in 2005 enabled individuals and businesses to build applications and services around location data [4, 5]. There are now 2.71 billion smartphone users worldwide – a number that almost doubled during the last five years [22]. Smartphones play an increasingly important role in people's everyday life, and so do the data produced by them for businesses. Sensors such as GPS, accelerometers, ambient light sensors, and compasses have become standard features of almost any smartphone [7].

Mobile computing and big data both have been flourishing fields of research on their own in recent time, however only a few studies explore the convergence of both domains [1]. Analyzing data collected by mobile devices enables businesses to gain novel insights form hitherto unrecognized sources [1]. This explains the existence of the relatively new research field of Mobile Big Data. Unlike "general" big data, mobile big data is distinct in the multi-dimensional, individual, personal, multi-sensory, and real-time nature of the data produced and processed.

2.2. Digital Data Streams

In recent past, the velocity of the data production

has rapidly increased, as computers, smartphones and tablets became constant companions of people. This enabled the capturing of everyday activities digitally, leading to a continuous flow of real-time data. This technological phenomenon is coined *digital data streaming* (DDS) [18].

For analytics purposes, large parts of mobile-generated data are DDS. Unlike *static data* that is stored in a database before it is processed, streaming data is processed as a so-called event-before-storage. Events are individual records consisting of field- and meta-data. Field data is the actual information contained of an event to be analyzed. The meta-data within an event determines how an event that is being processed by a certain application interacts with and is related to other events. If each event is processed as it is delivered and only the results are stored in databases, then the data is streaming [20].

Pigni, Piccoli, and Watson [18] investigate characteristics of DDS. According to them, a DDS can be human-generated (e.g. posts or tweets) or machine-generated via sensors (e.g. GPS location). An event of a DDS can contain up to six basic informational components: the time (when), the location (where), the actor (who), the activity (what), the tool that created the event (how), and the occurrence' reason (why). In this paper, we are mostly interested in the time, location, and the actor.

2.3. Customer Profiling in Financial Services

Prior studies such as Hormozi and Giles [12] and Shaw, Subramaniam, Woo, and Welge [21] investigate the early attempts of banks to engage in data mining. Data mining refers to any activity aimed to extract important information from existing data to enhance decision-making. Declining cost of data storage and computational power and increasing ease of collecting data in an increasingly digitally permeated environment are major trends that drive data-driven organizations to engage in data mining and predictive modelling [16].

Within the banking industry, one specific application field for predictive modelling is customer acquisition and retention. Revealing patterns in customer behavior enables advanced targeting, as it allows to promote products to a specific customer or group, which are expected to be interested in a given product or service. High cost for non-targeted, universal and often ineffective promotions can be cut [9].

Prior studies about data mining argue that conventional approaches to customer segmentation lead to increased cost and low response rates, as it is hardly possible to cluster customers into homogeneous groups [21]. Especially when selling financial products like

credits, highly heterogeneous preferences among customers are to be expected [8]. The field of knowledge-based marketing uses data mining to tackle these problems. Making use of large accumulations of customer preferences in databases, advanced customer profiles are created. Shaw et al. [21] claim that customer profiling relies on demographic consumer details as well as characteristics of past purchases, such as frequency, size or amount, and recency of a transaction. In a different but related study, Kitchens, Dobolyi, Li, and Abbasi [14] examine the field of advanced customer analytics. The authors argue that key to superior insights is the use of complementary data sources and their integration into models across functional silos inside and outside of an organization. However, transaction data and customer demographics only serve as the basis for a customer analytics solution.

A recent study by Martens, Provost, and Clark [15] introduce a model that employs a bank's customer transaction data to predict which consumers will be promising prospects for specific offers, pointing out that most target marketing initiatives of banks include information about the frequency, the time, and the amount of a certain transaction. The study refers to this data as structured data but argues that transaction data can also include fine-grained information on consumers' behavior. One way to generate fine-grained information is through network analyses. This is done by analyzing, which merchant customers are paying in a transaction. Customers paying to the same merchant are expected to be more similar in their behavior. The more connections they share, the more similar are their preferences. By including the data of the connections of customers in the same model, the authors show that the prediction output can be enhanced compared to a model that relies only on time, frequency, and amount of a transaction [12].

Another relevant stream of literature in this area pays attention to attrition rate and how to retain existing customers [13]. Hu [13] employs various methods (decision tree, boosted naïve Bayesian network, selective Bayesian network, neural network, and the ensemble of class of the above methods) to build clusters of customers to eventually use these for a more precise targeting. A credit card data warehouse served as the primary source of data for retention modeling. Additionally, third party data such as account related demographic and credit bureau information was used, supplemented by segmentation documentation and a payment database containing all processed checks of customers.

2.4. Location-based Financial Services

Banks have been known to exploit various sources

of data for analyzing and predicting customer needs [12, 13, 19, 15, 21]. These data include demographics (e.g. gender and age), past purchases (e.g. credits, bank accounts), payments (e.g., frequency, amount, and time of a transaction) as well as meta-data generated by models through transaction data [15].

The assessment of recent developments in technology has revealed that the increased use of mobile devices leads to the generation of novel data streams with unique characteristics [6, 7, 18]. One such data stream is location data captured by GPS sensory components in smartphones. The above assessment of literature and our anecdotal evidence of the European banking sector unveiled that banks have not yet embraced the value of adding real-time location data in their sales approaches and models. Many banks are still exploring internal machine learning use cases, trying to integrate their disparate data silos by setting up a shared data warehouse and data lake.

We argue that adding a customer's real-time location data to profiling models can significantly enhance customer profiling and prediction in offering financial services and at the same time reduce cost. Real-time location data is not new to all financial services providers, as some use it for fraud prevention. Examples include credit card firms Mastercard, Visa, and American Express which block transactions when they happen in close temporal proximity in disparate geographical regions. Bonus plan provider Payback uses location data to enable participating retailers to market their products to individuals near their stores, further corroborating that there is a market for such services that also banks can tap into by virtue of their business model. We intend to embrace real-time location data in form of a spatio-temporal component to prediction models.

3. Methodology

In line with design science research (DSR) the primary goal of this paper is to develop a model that efficiently predicts the wishes and wants of a customer in each environment [10, 11]. The development of such a model seamlessly aligns with the creation of a novel artifact described in DSR [17]. The literature review has shown that there is an identifiable business value in having such an artifact. In DSR, the practical utility embodies a central role, ensuring research relevance [10, 3]. On the other side, DSR employs the academic knowledge base of existing theories, models, and frameworks to build the artifact. Designing the artifact allows not only to contribute and add knowledge to the associated research domains – in this case mobile data exploration, customer targeting, etc. – but to create immediately visible benefits to practitioners.

The search process for an efficient artifact solution to a problem is firmly anchored in the DSR approach. We conducted five (5) interviews and reviewed archival documents to help us establish the objectives that our artifact would be fulfilling. First, we conducted 2 interviews with senior executives of a major German consulting company that specializes in advising financial services clients and 3 interviews with representatives of client banks. Second, we reviewed 18 consulting and industry reports. This is of importance for solutions like this, which apply relatively new methods (e.g., location data and data analytics) to a given the problem. Often, several attempts to find a suitable solution are needed. DSR provides the liberty to generate design alternatives and test them against previously defined goals [11, 17].

4. Artifact Description

4.1. The Process

Prior to discussing the technical implementations of the artifact, we provide an overview of the designed artifact by introducing a process model shown in Figure 1. On a higher level, the model can be divided into two phases: The location model and the decision model. Simplified, the process is as follows:

a) The customer shares his/her geographical live location
b) The location model determines if a store is close to the customer and if the store is important to him/her
c) If yes, this information is passed on to the decision model that determines if a customer is likely in need of a credit
d) If yes, an offer is sent to the customer and the feedback (whether it is accepted or not) is used to update the decision model

The starting point is the customer who transmits live location data through the mobile device. This location data is matched with the location of stores of potential interest to the customer. These stores of potential interest are determined upon the historic account transactions. Stores that appear as a creditor in the client's account data on a frequent basis are presumed to be of interest to the customer. Once a customer is close to such a store a match occurs which will be referred to as a location event. For each location event, information about the location, date and time, and account balance of the customer is captured. This information is forwarded to the decision model which predicts the probability of a customer to be
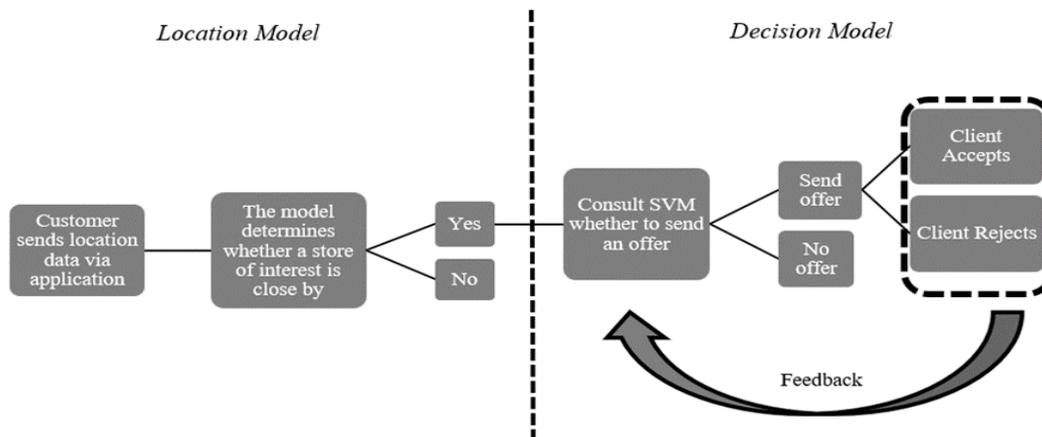
Figure 1. Process Overview of the Model

likely to accept a credit for a location event. A classification model is initially trained with a dataset generated through historic account transaction. Once this model predicts, that an offer will be accepted by the client, it is sent out. The client can either accept or reject the offer. After the client has made a choice, this observation is stored and added to the training data of the model. Thus, training data is continually updated, and each following prediction is made upon all previous choices by the client.

4.2. The Location Model

The model uses live location data of customers and combines it with historic account information. Two data sources feed the model: historic transactions of a customer and live location data. To train the model, a dataset containing anonymized historic transaction and customer data was provided by a leading German provider of multibank-capable banking solutions. The dataset contained all bank account transactions of a customer from the past two years, totaling 679 transactions. The second data source is live location data which is ingested into the model in form of latitude and longitude data as well as a time stamp.

Since one of the paramount intentions of this artifact is to obtain supplementary value from location data when modelling a customer's preferences, the initial stage of the model is concerned with the creation of an interface for constant retrieval of live location information. The sole retrieval of live location data is of little value until brought into context and combined with other data. To build the desired artifact, there is a need to gather insights on the specific stores the customer visits frequently, and thus, are of interest. Subsequently, the live location data can be used to determine, whether a customer is close to such a store. To retain information about stores of potential interest, the model starts by processing information contained

in the account data set. For this purpose, the creditor of a transaction account data delivers valuable content, as it specifies the payee. In the first stage, this attribute is filtered, so that only in-store purchases appear. This is a necessary step, since recurring transfers such as rental payments, insurance payments, gas or electricity bills are not of interest for this model.

Upon this data, another dataset is generated, which provides information about the frequency of appearances of each creditor in the account transactions file. Each creditor is ranked upon the number of transactions the customer has made within the previous two years. This dataset offers an overview of the most frequently visited stores and, based on that, allows to define stores of interest to the client. For the implementation of this model, a threshold was set, which defines stores which were visited more than ten times in the last two years as stores of interest to the client. This threshold is not fix but can vary as desired by future users of the model.

Once the stores of interest are determined, the next step is concerned with the acquisition of their location data in order to later match these with the live location data of the customer. To do so, a geographical region is defined. For the implementation of our model the City of Cologne, Germany was chosen. The purpose of this stage of the model is to create a dataset containing all branches of stores of interest to the customer within the given region. If a customer, for instance, frequently shopped at Shell gas stations, all Shell gas stations within the given geographical region need to appear in the dataset. We collected this information from Google Maps and Open StreetMap and created a dataset that contained all branches of a store of interest with addresses within the defined geographic region.

To ease a later match with the live location data, we rounded the latitude and longitude of each store to three decimals which corresponds to an approximate accuracy of 100m. To determine a location with one

single argument instead of two (longitude and latitude), the rounded latitude and longitude are concatenated and form a unique location identification, which is referred to as a location code.

Each store's location code defines its location with only one argument and with the proximity of 100m. Based on this location code, the live location data will be matched to determine if a customer is close to one of his or her preferred stores. Hence, the live location data is formatted in the same way and a location code is created for each live location observation that enters the model.

This constitutes the necessary preprocessing of the data to enable a matching between a customer's stores of interest and the live location. Every time a customer sends live location data, this is being transformed into a location code which is then compared to the location code of any of his stores of interest. If a customer is within the range of approx. 100m to one of these stores, the model generates what is further labeled a location event. However, if a customer's live location data does not match any of his store's location, no event occurs. Hence, a location event is the necessary foundation to initiate further processing steps of the model. For each location event, the following information was saved in a dataset: store, date, time, and current account balance.

4.3. The Decision Model

Once a location event occurred, this information is passed on to the second stage of the artifact: the decision model. For each location event that takes place, the decision model makes a prediction, whether the customer is likely to accept a credit offer at this specific store at the given time. As soon as the probability of accepting a credit is predicted to be higher than .7, an offer is sent, otherwise no offer is sent. If more than one location is in reach, the store with the highest probability is picked. In other words, the decision model attempts to classify each location event into two classes: accept or reject. To perform this task, a support vector machine (SVM) is used. The SVM partitions a given population of individuals Ω into two classes $\{-1,+1\}$. For each individual i $\in \Omega$, there is:

$x_i = (x_{i1}, x_{i2}, ..., x_{ik})$: vector of explanatory variables
$y_i \in \{-1, +1\}$: class membership.

The goal is to predict the class membership of new individuals for which the explanatory variables are known, i.e.,

$$x_{new} \rightarrow y_{new} \in \{-1, +1\}$$

An SVM defines a hyperplane ($w^T\mathbf{x} + b = 0$), which is a boundary in a n-dimensional space, that best separates the two classes. The hyperplane acts as a discriminative classifier and is later in the prediction phase used to classify new objects depending on their positioning in the multi-dimensional space:

$$y_{new} = \begin{cases} +1 & if \ \mathbf{w}^T \mathbf{x}_{new} + b > 0 \\ -1 & else \end{cases}$$

To build such a model and to find an adequate decision boundary, data is needed. Data that initially serves to define the decision boundary is referred to as training data. This training data, i.e. observations where a customer accepted a credit and observations where a customer rejected a credit in the past, is created by using the account transactions of the last two years of the client. For each transaction, a given class membership accepted or rejected is defined. To do this in a meaningful way, an economic assumption is made. It is assumed that the interest payable for a negative account balance is higher than the interest payable for a consumer credit. Thus, an economically acting consumer will prefer paying interest for a credit over overdrawing the account. For every transaction that took place in the past, the amount of the transaction is compared to the balance of the bank account at the specific time of the transaction. For all cases where the account balance was smaller than the amount spent in a transaction, it is assumed that a value maximizing customer would accept a credit because its interests are cheaper than the interest due to overdrawing the account. Using this method, all account transactions were classified either as "accepted" or "rejected". This data serves as training data to set up the SVM and to initially define a decision boundary between the two classes.

For our model, the explanatory variables are the store, date, time and account balance. This information is contained in every account transaction but can also be gathered for each location event. Ultimately, the target is to predict whether a customer accepts or rejects a credit offer based on what his or her account balance at a given date and time at a given location (i.e., store).

To employ the model in a meaningful manner, however, the data needs to be further aggregated. The model converts given dates into weekdays. Having the individual weekdays on which a customer shops or solely appears at a certain store allows to discover more sophisticated patterns on buying behavior and preferences than just a simple date. Further, the store is used as an explanatory variable. We encode them as one-hot encoded variables. All previously defined stores of interest which appear as creditors within the transaction data are used as explanatory variables. The balance of the bank account and the given time is also added to the information.

Each explanatory variable represents one separate dimension in the multi-dimensional space. Stores and weekdays are binary encoded, while the time and the account balance are expressed as integers. The class membership ("accepted") is represented as a Boolean which can take the values of "true" and "false". Having transformed the data, the SVM is trained upon the classified transaction information. The classification achieved with the training of the SVM is now used to classify location events. Each location event enters the model and is predicted to be in one of the classes.

As described in the location model section of this paper, the location event first determines a specific branch of a store chain by matching its location code with the live location code of the customer. However, not the specific location code or address of the specific branch is added to the information, but the name of the store chain. The same information is provided in each account transaction, where the store chain – and not a specific branch – appears in the creditor column. Thus, stores in location events have the exact same title as stores in the account transactions, which allows a direct match. This abstraction from the individual branch enables to make statements about the general preferences regarding a chain of stores. Consequently, patterns can be detected across different branches. If the model determines, that a customer is likely to accept a credit for a given store chain at a given time, this will apply to all branches of this store in the defined region.

It is important that the explanatory variables are in the exact same format for both the initial account transactions and the location events to enable a prediction and a later accumulation of instances for training. As the data of a location event is prepared, it enters the SVM for prediction. Depending on the outcome of the prediction, the information contained in the location event is subsequently used as further training data to improve the model.

The SVM model predicts a probability for a class membership for the location event. If the probability, that the new instance is within the "Accepted" class is higher than .7, an offer is sent to the customer. Otherwise, the information is not further processed, and the client receives no offer. For each new instance (location event) for which the model has predicted to send an offer, an entry in the training data set is added, leaving the class membership blank because the feedback from the customer to accept or reject a credit offer is pending. Once the customer made a choice, this information is added to the training data to complete the entry. Every time, a new instance is added, the model is trained again with all available training data up to that point. Thus, all upcoming predictions of the model include the newly created instance, which allows the model to continuously learn from the provided customer feedback in a transparent and understandable manner.

By continuously feeding customer feedback into the training data of the model, the number of entries increases over time. As a result, the originally generated training data will play an increasingly subordinate role as the ratio of new information to the original information increases. This is intentional, as the actual interests of the customer are consequently increasingly considered. The initially generated training data from the account transactions, hence, solely serves as a basis to start the model and make first logical predictions based on economic assumptions that are basically common-sense.

5. Evaluation

This section seeks to demonstrate how the artifact functions and fulfills the previously defined goals. The achievement of goals in areas outside the development environment are assessed [10]. Table 1 provides an overview of how each part of the model contributes to providing a solution to the problem.

First, it will be evaluated, to what extent targeted offers are considered in the model. An element associated with these are the stores of interest. Therefore, the account data plays an important role, as it allows to continuously identify the most recently visited stores. Because the account data are filtered by the shops that the customer has visited very often lately, targeting is already achieved at one the initial steps of the model. Less frequently visited businesses that are of minor interest to the client are therefore excluded at this first step. Although the general preferences of a customer are identified by the targeting, there is still the possibility to create an overflow of targeted offers. Consequently, another objective is to not spam the customer. This is accomplished in two steps by the model: First, the location model fixes a location match when a customer is nearby a store of interest. For all situations, a customer is not close to one such store, the model does not send an offer. Two datasets of live location data were imported into the model. One record contained coordinates that did not match any of the stores of interest coordinates. Another record contained coordinates describing a point near one of the stores of interest. Both records were entered as variables. For the first record, no location event took place, and the model outputs an empty dataset. For the second record, in contrast, the model successfully matched the live location to one of the stores and wrote this information in a data frame and created a location match.

Table 1.Contributions to the Problem Space

	No Spam	Targeted Offers	Learn from Feedback
Location Model	Depending on the physical location of a customer, the sending of an offer is initiated or restricted.	Only stores that appeared more than x times in the last y months in the account data are considered.	
Decision Model	Based on whether a customer has previously accepted credits in similar situations, the sending of an offer is initiated or restricted.		Offers that are sent to customers serve as further training data, as soon as a feedback is received.

In addition, the model employs a more precise filtering of offers in the second stage – the decision model. Based on the results of the support vector machine, offers are sent only for certain location events. To demonstrate this capability, following simulation was performed: To test the prediction, we again assumed that a customer always accepts a credit if the amount of the purchase is greater than her bank account balance, though in reality not all customers will make their choice in accordance with the above stated assumption. When implemented, customers will determine their choices upon individual motives. These choices are reflected in the feedback of a client which serves as further training data. In contrast to the initial training data, which is generated upon the above stated general assumption, this data is individual training data, reflecting the individual preferences of a customer. To demonstrate this effect, three customer profiles were artificially generated. Subsequently, the model was trained upon these profiles. For each customer profile, specifically edited training data were created according to some assumed preferences. Each training data set contains the initial training data and, in addition, 400 instances through which the assumed preference is mirrored. For 200 of them, the assumed preference applies, and they were tagged as "accepted". For the other 200, the assumed preference does not apply. These instances were tagged as "rejected". For each profile, two test scenarios were predicted. One, where the above defined condition was fulfilled, and another, where it was not fulfilled. In addition, the initial model – without the individual training data – was used to predict each given scenario. Thus, it can be tested, whether the model adopts to each customer profile. Comparing the prediction of the model trained upon the initial data with the individually trained model allows to determine, whether the model is capable to include the added individual training data in its decision base and is thus learning from a customer's feedback.

While the model which was trained upon the initial training data did not correctly predict the customers preference in all times, all predictions of the model which is trained upon the individual profiles led to the correct classification. This proves that the model incorporates the feedback from a costumer, includes it into the training data, and ultimately makes superior predictions as it learns from it.

Moreover, the accomplishment of these objectives has several effects on forecasting consumer demands in general and for consumer credits. First, the model provides an inexpensive, yet, accurate way to anticipate customer needs. With the help of the selected classification method one can now very efficiently respond to the individual needs of the customers. This equips banks with the necessary information to confidentially approach a client with a personalized offer, at the right time and place. This allows a fundamental change in customer-bank relationships. In the past, the customer-bank interaction was mostly customer driven as it involved a certain need, such as a bank transfer or loan request. Having superior insights into customer desires provided by the artifact, allows banks to move away from their rather defensive position within the customer-bank relationship and focus on meaningfully offer time and place specific, targeted products.

The artifact can be used as part of a bank's strategy to expand their services portfolio into the mobile channel. The mobile channel is gaining prominence and many established channels of customer interaction like local branches are becoming obsolete, forcing banks to adapt to changing customer behaviors in order to remain relevant in the future.

7. Discussion

Gregor and Hevner [10] present a knowledge contribution framework for design science research. The

framework clusters contributions according to the maturity of both the application domain and the solution. We argue that the proposed artifact provides a novel solution to a known problem [3]. The literature research reveals no evidence that banks have yet employed location-based customer prediction systems to this date. However, the desire to predict a customer's preference is not new, as previous attempts to cluster customers into segments and target them accordingly are known beyond the marketing literature. Thus, the artifact is an improvement, as it is a new approach to solving a known problem. The developed artifact is unique in multiple dimensions and differs from past research attempting to predict consumer needs.

One such dimension is the general level of abstraction. Many other models which attempt to predict customer preferences are based on segmentation approaches. Hereby, customers are grouped into clusters of peers with similar interests and preferences. Upon this, each group is presumed certain preferences and customers who are members of a certain group receive according offers [21]. Our model instead assumes a "segmentation of one", i.e., customer preferences are predicted for an individual customer by evaluating only his or her individual information. The trade-off between generalizability and precise targeting is increasingly fading away, as the cost of precise, individual targeting decreases. This artifact takes this development into account and applies statistical techniques on an individual basis.

Additionally, and most importantly, the artifact incorporates another dimension – spatio-temporal information – by adding location data to a customer preference prediction model. As the increase of smartphones is on the rise, this novel data stream will not only create value in the context of predicting the need for loans, but also in other fields. Knowing where a customer is at a given time and take that information into account in real-time can add tremendous value to multiple kinds of analytics and prediction models. This artifact is one of many possibilities to incorporate location data for precise prediction of customer needs and illustrates just one use-case.

Depending on the specific context in which the model is applied, some challenges might arise during implementation. One of such challenges is scalability. While this artifact shows the concept based on one customer profile, practitioners need to have a certain infrastructure and toolset to adopt the model in a large scale.

When implemented, the model needs to be trained for each individual customer. Once done, it is necessary for the practitioner to test the model's prediction accuracy. As the predictions of the model constantly change as new observations are added to the training data, keeping track of whether the accuracy increases, is essential. One approach to do so is to take a certain amount of random observations of the existing training data and extract these at a certain point. These observations are then referred to as test data. For each observation in the test data, the class membership will be neglected, and a prediction for them is made by the model upon the remaining training data. The outcome of this prediction is then compared to the actual class membership, which allows to determine the model's accuracy by comparing the proportion of correctly predicted observations with false predictions.

The artifact represents many ways in which mobile location data of customers can be used by linking it to traditional data, such as in this case existing account data, to gain deeper insights in customer behavior. Hence, further research can draw upon this approach and it can be concluded how the use of spatio-temporal data also finds application in other areas.

8. Conclusion

This contribution presents an algorithmic model which predicts the need of consumer credits by considering a given spatio-temporal context of the customer. To do so, two data types – historic account transactions and live location data – constitute the decision basis of this model. The artifact is split into two sub-models, namely the location model and the decision model. The first part, i.e. the location model matches the live location data of a customer with a dataset of beforehand defined stores of interest to the specific customer. These stores are chosen upon the shopping habits of a customer which are derived from the transaction history of the corresponding bank account. Once a customer is geographically close to such a store, the model creates a so-called location event, containing information about the specific store visited, weekday, time, and account balance. Subsequently, this information enters the second part of the model – the prediction model – which predicts the probability of the client accepting a credit offer. This prediction is made by a Support Vector Machine, which is trained upon previous observations of that specific customer. This training data is created of historic transactions of the client. Once the model decides that it is likely that a client accepts an offer, it is sent. Following, the client has the chance to either accept or reject the offer. The feedback of the client is afterwards injected to the training data, which allows the model to continuously increase the precision of predictions over time.

The artifact sets itself apart from existing solutions for consumer preference forecasting in several regards. First, predictions are made on an individual cus-

tomer level and are not derived from customer clustering approaches. Second, the artifact takes the spatial and temporal situation of customer into account. This makes it possible to submit situation-dependent offers for future implementers.

We do however acknowledge that while our contribution unveils what is possible by unpacking the affordances of mobile technology – specifically location-based data, there are still hurdles to be crossed for practical everyday utility of such a solution. These include technical issues – such as privacy concerns, ethical/social issues – preventing that vulnerable customers get caught in debt traps and legal issues – articulation of appropriate regulatory framework within which such financial service innovations can thrive [5]. We hereby present these as areas for future scholarship to explore in order to uncover the theoretical and practical guidance to the users, developers and regulators of location based digital mobile services in the financial sector.

In conclusion, our study provides an illustration of how location-based artifacts can be leveraged in creation and capture of value in financial services. From the perspective of the banks, we showcase how mobile technology can provide an additional avenue for them to potentially provide solutions to high impact areas of their customers everyday life as they fend off competition for FinTechs and other digital disruptions.

9. References

[1] Ahmed, E., I. Yaqoob, I. Abaker, et al., "Recent advances and challenges in mobile big data", IEEE Communications (February), 2018.

[2] Alt, R., R. Beck, and M.T. Smits, "FinTech and the transformation of the financial industry", Electronic Markets, 2018, pp. 235–243.

[3] Baskerville R, Baiyere A, Gregor S, Hevner A, Rossi M. Design science research contributions: finding a balance between artifact and theory. Journal of the Association for Information Systems. 2018;19(5):358-76.

[4] Bellavista, P., A. Küpper, and S. Helal, "Standards and Emerging Technologies Location-Based Services : Back to the Future", IEEE Computer Society, 2012, pp. 85–89.

[5] Carlsson C, Walden P. Further Quests for Value Added Products and Services in Mobile Commerce. ECIS 2002 proceedings. 2002 Jan 1:19.

[6] Cheng, X., L. Fang, X. Hong, and L. Yang, "Exploiting Mobile Big Data : Sources , Features , and Applications", IEEE Network 31(February), 2017, pp. 72–79.

[7] Cheng, X., L. Fang, L. Yang, and S. Cui, Mobile Big Data, 2018.

[8] Dahlberg T, Mallat N, Saarinen T, Tuunainen V. Efficient Service Production and Consumer Choice of Delivery Channels. BLED 2001 Proceedings. 2001 Dec 31:20.

[9] Decker, P., "Data Mining's Hidden Dangers", Banking Strategies (March/April 1998), 1998, pp. 6–14.

[10] Gregor, S., and A. 8, "Positioning and Presenting Design Science Research for Maximum Impact", MIS Quarterly (June), 2013.

[11] Hevner, A.R., S.T. March, J. Park, and S. Ram, "Design Science In Information Systems Research", MIS Quarterly 28(1), 2004, pp. 75–105.

[12] Hormozi, A.M., and S. Giles, "Data Mining: A ompetitive Weapon For Banking And Retail Industries", Information Systems Management, 2001, pp. 62–72.

[13] Hu, X., "A Data Mining Approach for Retailin Bank Customer Attrition Analysis", Applied Intelligence, 2005, pp. 47–60.

[14] Kitchens, B., D. Dobolyi, J. Li, and A. Abbasi, "Advanced Customer Analytics: Strategic Value Through Integration of Relationship-Oriented Big Data", Journal of Management Information Systems 35(2), 2018, pp. 540–574.

[15] Martens, D., F. Provost, and J. Clark, "Mining Massive Fine-Grained Data To Improve Predictive Analytics", MIS Quarterly 40(4), 2016, pp. 869–888.

[16] Mitchell, T.M., "Machine Learning and Data Mining Over the past", Communications of the ACM 42(11), 1999.

[17] Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. Journal of management information systems. 2007 Dec 1;24(3):45-77.

[18] Pigni, F., G. Piccoli, and R. Watson, "Digital Data Streams", California Management Review 58(3), 2016.

[19] Poel, D. Van Den, and B. Larivi, "Customer attrition analysis for financial services using proportional hazard models", European Journal of Operational Research 157, 2004, pp. 196–217.

[20] SAS Institute, "Streaming Data vs. Static Data", 2018. http://documentation.sas.com/?docsetId=esptexanddocset-Target=streaming-static.htmanddocsetVersion=4.3and-locale=en

[21] Shaw, M.J., C. Subramaniam, G. Woo, and M.E. Welge, "Knowledge management and data mining for marketing", Decision Support Systems, 2001, pp. 127–137.

[22] Statista, "Number of smartphone users worldwide from 2014 to 2020 (in billions)", 2019. https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/