Zeppelin Universität

Department Corporate Management & Economics

Dr. Manfred Bischoff Institute for Innovation Management of EADS |
Assistant Professor for Innovation and Entrepreneurship, in particular
Innovation Communication and Behavioural Research

Jun.-Prof. Dr. rer pol Marco Hubert

Master's Thesis

The Chances of Data Mining in Subscription-based Businesses: A Literature-based Analysis

Author:	Kerstin C. Weinheimer
Number of Matriculation:	11200550
Course of Studies:	Corporate Management & Economics (MA nk)
Semester:	Fall Semester 2013
Supervisor:	Jun.-Prof. Dr. Marco Hubert
2nd Supervisor:	Dr. Marc Linzmajer, Uni St. Gallen
Date of Submission:	06 December 2013

Abstract

This paper analyses the chances of data mining for subscription-based businesses.

The key topic in this context is churn management. This paper argues that only the identification of customers with high probability to withdraw is no advantage in itself. Successful retention is crucial instead. In order to focus on customers with maximum margins, the concept of Customer Lifetime Value is applied. This leads to the development of a descriptive framework. This framework quantifies the returns of data mining and thus shows the beneficial effects of applying data mining methods (churn prediction- and CLV-models) to consumer data by summarising the retained profit.

The paper also makes proof of several research gaps. Finally, it suggests the development of a predictive framework based on the one developed for measuring the returns of data mining using retention probability.

Table of Contents

List of Tables

List of Figures

"It's about getting things down to one number.
Using the stats the way we read them,
we'll find value in players that no one else can see. [...]
Like an island of misfit toys."

Peter Brand to Billy Beane in:
Moneyball (US 2011, D: Bennett Miller)
(Min. 27-28).

1. Introduction

In March 2013, Amazon announced the acquisition of Goodreads (cf. Amazon.com, Inc. March 28, 2013). But what exactly is Goodreads? Goodreads is a social reading site that "has over 20 million members who have added more than 600 million books to their shelves and written more than 25 million reviews." (Goodreads 2013). It provides people with a platform to subscribe, to share what they enjoyed reading and what they are planning to read, creating a valuable book discovery engine for new books.

The interesting question now is - why did Amazon pay an assumed price of 150 million dollars (cf. Thomas, March 29, 2013; Carmody, March 29, 2013) for a company that does not have a turnover generating business model? Using the words of a TechChrunch article, "[t]he amount of data that Goodreads has on its users alone makes the acquisition a slam dunk." (Olanoff March 28, 2013) According to this statement, data must have been the asset Amazon was willing to pay for amongst other benefits. Hence the question is arising – what makes this data so valuable? What is the potential of these amounts of probably rather unstructured information that justifies nine-digit expenses? This is one of the key questions this work is going to investigate.

The question for the potentials of big data is especially interesting considering the fact that not every company has the same investment capabilities like Amazon, and not every company is yet willing to make the necessary investments into data mining.

Subscription-based businesses provide particularly interesting data that is well suitable for exploration because of the panel data they can provide, and because of the possibility to assign actions to individuals. Moreover, subscription-based businesses have a particular interest in the chances of data mining: the possibility to predict customer churn. Thus, over the last five to ten years, this has become a central concern of research working with data mining. This is one of the reasons why the analysis focuses on data mining in the context of subscription-based businesses. However, this is going to be explained in more detail in section 2.

With regards to the structure of this work, the remainder of the paper is organised as follows.

Section 2 sets the foundation by first outlining the relevance of subscription-based businesses, ranging from telecommunications to media subscriptions. It defines the term and shows what motivational factors for subscribing exist on the consumer side. Secondly, the relevance and meaning of data mining are explained. The latter is especially important, since data mining has become a much hyped phrase and its frequent use might have blurred its concrete meaning. As this analysis is, as a whole, aiming to portray not only the current state, but also to unravel research opportunities, fundamental definitions of this work need to be clear. Therefore, the section starts highlighting the relevance of the topic, outlining the context and introducing the subject.

Section 3 focuses on the chances of data mining in subscription-based businesses. The main 'chance' of which is churn management, essentially. Hence, different chapters define churn, give a categorisation of churners, show what a churn prediction model is and which data is useful to build one, and round it off with a comment on predictive accuracy and interpretability.

Section 4 then introduces the analysis of current and relevant literature on churn management. It analyses the current state of research, shows prevailing trends – the concentrations on predictive techniques and the telecommunication industry - and consequently highlights research gaps. These research suggestions introduce the next section.

Section 5 presents a framework based on the findings made so far. The chapter argues that only knowing about a customer's intent to retreat is not an advantage in itself, but that organisations need to act upon that knowledge - with retention measures. Ideally, those efforts should focus on the most profitable customers. Therefore, section 5 will introduce the concept of Customer Lifetime Value, in order to be able to determine those promising customers. In the end, this framework will result in a metric – the summarised, secured profit – that quantifies the ‘chances of data mining’. This method for quantification is the key result of this work.

Because the framework is built of several newly introduced parts, section 5 encompasses chapters on Customer Lifetime Value, Customer Relationship Management, retention measures, direct marketing, switching costs and retention probability. In order to avoid ambiguity, a chapter on Customer Lifetime Value models selects the most appropriate estimative approach. In the course of this analysis, another literature overview shows the lack of applying this profit-oriented approach to subscription research.

Section 6 suggests the development of a predictive model based on the descriptive framework developed in the previous section as a future research task. For reasons of security and transparency it might be of interest for organisations to know about the possible gains of investments in data mining in advance. Key to this model is the determination of retention probability, which was introduced previously.

Section 7 summarises and concludes. It summarises the findings made, recapitulates this work’s research approach and discusses not only the findings made based on the literature analyses, but also the concept developed. It summarises the research gaps and outlines three directions for future research.

In conclusion, this work shows the chances inherent to data mining for subscription-based businesses. Consequently, it developed a framework to quantify them, creating a new metric; all on the basis of the analysed literature.

2. Subscription-based Businesses and Data Mining

What is so special about subscription-based businesses? Not only have they invaded many areas of daily life, such as telecommunications, newspapers, or media content e-commerce (e.g. Spotify or Netflix) but they also play a considerable role in economy. Their popularity might be due to the fact that the business model ensures stable turnover that can be relied on. Moreover, they dominate the field of research exploring the chances of data mining. The reasons behind this are various, but one is central: it is the economic interest that lies behind what the exploitation of data mining abilities can bring to subscription businesses: the security of stable numbers of consumers.

Due to increasingly “saturated markets and intensive competition, a lot of companies do realize that their existing database is their most valuable asset.” (Coussement & Van den Poel 2008: 313). This is a trend particularly intense in subscription services, because it is much “more profitable to keep and satisfy existing customers than to constantly attract new customers who are characterized by a high attrition rate” (Coussement & Van den Poel 2008: 313). This is an argument also put forward by Hadden et al., saying that it is more costly to “win new customers than to retain existing ones.” (Hadden et al. 2005: 2902). Therefore, it can be said that it is the argument advocated by most of the articles on churn management in the subscription context (cf. Hung, Yen & Wang 2006; Verbeke et al. 2011; Verbeke et al. 2012; Kim & Moon 2012; Kim, Lee & Johnson 2013). This circumstance even persuaded Karahoca and Karahoca to write that “[i]t is well known that the cost of a retaining a subscriber is much cheaper than gaining a new one from another” (Karahoca & Karahoca 2011: 1814).

Some studies go so far as to publish estimations of how much it can cost a company to replace a customer: “the costs of acquisition is typically five times higher than retention.” (Hung, Yen & Wang 2006: 515) This is a commonly acknowledged number based on studies conducted in the 1990s (cf. Verbeke et al. 2012: 211). Lemmens and Croux, even present the concrete amount of “$300–$700” (Lemmens & Croux 2006: 277). The general validity of those numbers can probably be doubted, because of variations in subscription price, retention incentives, marketing expenditures and industry. However, the fundamental tenor is the same: it costs the company to replace a customer lost. “Because the cost of replacement of a lost wireless customer amounts [...] in terms of sales support, marketing, advertising, and commissions, churn may have damaging consequences for the financial wealth of companies." (Lemmens & Croux 2006: 277) Besides, when the market is highly saturated, it might be not only more expensive but even impossible to attract new customers.

A second argument in favour of avoiding customer churn is the circumstance that "on the revenue side, studies have shown that continuing customers purchase higher volumes at higher margins" (Keaveney & Parthasarathy 2001: 375). This was confirmed and expanded by Verbeke et al.: "long-term customers generate higher profits, tend to be less sensitive to competitive marketing activities, become less costly to serve, and may provide new referrals through positive word of mouth" (Verbeke et al. 2012: 211). The beneficial monetary effect is also advocated by Kim and Moon, who claim that with successful means of retention, customer profitability can be strongly increased (cf. Kim & Moon 2012: 11718). Consequently, preventing churn not only secures existing turnover, but it also prepares the way for an increase in future turnover. Moreover, Keaveney and Parthasarathy indicated that retained customers even increase their usage of a service even though when prices rise. To sum it up, it becomes clear that continuing customers are very lucrative for a company and have to be treated accordingly. As it seems, they may actually be preferred over newly acquired subscribers.

To sum it up, there is a real economic interest lying behind the research on models predicting customer cancellation. The interest is especially intense in markets that are becoming progressively saturated, such as the mobile communications market, predominantly. This refers to the markets in Asia (cf. for example Kim, Park & Jeong 2004), but also to the American and European market: "As with many other sectors (e.g., the newspaper business), churn is an important issue for both the U.S. and the European wireless telecommunications industry." (Lemmens & Croux 2006: 277)

Unfortunately, the telecommunications industry is struck with abnormally high churn rates: "[i]n particular, with exceptionally high annual churn rates (20-40%), the firms in mobile telecommunications industry want to develop predictive models that accurately identify which customers are most likely to terminate the current relationship" (Kim & Moon 2012: 11718, cf. also Lemmens & Croux 2006: 277). This is a view that has been gaining weight in research for several years. Meanwhile, also "companies have acknowledged that their business strategies should focus on identifying those customers who are likely to turn." (Hadden et al. 2005: 2902)

The development of data mining techniques has of course supported this trend of churn management, because "[t]he explosive growth in databases has created a need to develop technologies that use information and knowledge intelligently." (Liao, Chu & Hsiao 2012: 11303)

These are the reasons for the amount of research done on churn management of subscriptions. In some articles, it has thus been commented that "estimating the churn probability has been most actively and rigorously studied" (Kim & Moon 2012: 11719). Those are also the reasons why this work is going to focus on the topic of subscription-based businesses offered B2C. Moreover, the topicality of the issue, which caused great results from the academic sphere, provides an ideal base for the literature analysis in section 4.

In this context it is necessary to explain the reasons for excluding subscriptions running on a B2B basis. There is a minor field of research on database subscriptions of libraries (cf. e.g. Thohira, Chambers & Sprague 2010), which is excluded. This field has to be omitted for two reasons. Firstly, libraries are institutional customers and not individual ones, and secondly, the studies were conducted from libraries' perspectives and not from a neutral third party researcher's point of view.

Before dealing with the analysis of churn management itself, however, two terms have to be set clear, preparing common ground for the upcoming chapters and conceptual developments. Those terms are 'subscription' and 'data mining'.

2.1. Subscription-based Businesses

In their 2011 work, Samimi and Aghaie give a useful and precise description of the term ‘subscription’ following Dover and Murthi (2006). They distinguish companies offering products or services on a contractual basis, meaning subscription-based, and those offering pay per use:

> “Such corporations can operate on either contractual (i.e., subscription-based model) or non-contractual basis (i.e., pay per use) (Dover & Murthi, 2006). This paper investigates the contractual business model or the so-called subscription-based services. Customers of such a firm usually purchase a quarterly, an annual, or an “annual billed monthly” subscription.” (Samimi & Aghaie 2011: 89).

This definition shall serve as the basis for the following research, as it makes clear what is important for a subscription-based businesses. The provider has to offer a contract to obtain the right of usage for a product or service. This happens in exchange for regular payment. These regular intervals can range from every month up to only once a year.

The subscription-based model, of course, does not only appear in the telecommunications industry, which was mentioned earlier. Actually, it is of use in many different industries. As Samimi and Aghaie state, “a subscription-based service provider [...] may involve a broad spectrum of service industries such as telecommunications, insurance services, internet service providers, as well as many internet-based service firms like online content providers.” (ibid.) The latter ones are of special interest, since their number has strongly increased with the spread of e-commerce and the online availability of flat rate models (Shy 2008: 2413) for media content such as music, films and books.

In this context, some of the online content providers referred to above shall be introduced. For music, there are Spotify, Pandora, Napster, Simfy, Deezer, rdio or Grooveshark. For movies, the willing consumer can chose between Netflix, Lovefilm, Hulu and Watchever. Oyster, Legimi, Safari Books Online, Paper'C, Skoobe, Amazon Prime, 24symbols, Litfy, Total Boox and Scribd offer subscriptions for the engaged reader. As shows this number of content providers, it is a market currently massively developing and changing at considerable pace. In particular, the e-book-business is under constant development and numerous international start-ups are surfacing around the globe.

Read Petite, for example, was announced to launch at the end of 2013 (cf. Abrams April 14, 2013). Booxl, ‘another Spotify for ebooks’, was announced as well, but has not yet appeared (cf. dbw November 23, 2012). The same is valid for Huffier, which is still in beta (cf. Huffier 2013); versus Oyster, a start-up that was indeed launched as recently as in September 2013 (cf. Bertoni September 05, 2013). Oyster claims ‘to change publishing, the world of books and the way of reading’ (cf. Oysterbooks.com).

Their purely digital business models create unique opportunities for data mining and telling insights into customer behaviour. It is probably a promising future research opportunity to apply the data mining techniques that have been developed so far to their behavioural data.

Samimi and Aghaie have grasped those “rapid alterations of market conditions” (Samimi & Aghaie 2011: 89) and connected them with assumed changes in customers’ needs (cf. ibid.). Therefore, they argue, companies need to register sudden changes in customer behaviour – and to “have a prompt reaction against unusual shifts in customer usage pattern.” (cf. ibid.) They claim that such alterations in customer behaviour can be detected by exploring usage data (cf. ibid.). Besides churn management, this is another interesting but rather unexplored area of data mining applications: using behavioural data to detect unorthodox behaviour – and using it to adapt the business model, for example, or at least to develop and run “appropriate marketing campaigns and service customizations” (ibid.). However, this topic of adapting the business concepts cannot be investigated any further, as it lies outside of this work’s thematic focus.

So far, the importance of researching on subscription-based businesses has become clear. Moreover, the relevance of this issue was also advocated very recently by Fruchter and Sigué: “The importance of studying subscription services becomes self-evident in light of the fact that, at present, almost every household in a Western country is involved, in one way or another, in these services.” (Fruchter & Sigué 2013: 2180)

With regards to the aspect of regular payment that is one of the key characteristics of a subscription, subscriptions are usually organised as follows:

> “The four most common components of a subscription service price are: (1) an activation or installation fee, which is generally a one-time payment the customer makes on entering into an agreement with a service provider; (2) a subscription fee, which is an ongoing payment paid continually; (3) a usage fee paid for additional specified usages in addition to the basic service; and (4) an exit or cancellation fee paid for early termination of the subscription contract.” (Fruchter & Sigué 2013: 2180)

An interesting characteristic of the nature of subscription fees is the fact that customers stay longer with a company when they are being charged monthly instead of in larger intervals: “[t]he results show that customers who are billed monthly for annual subscriptions maintain their subscriptions longer than do customers billed quarterly, when compared to annual subscriptions.” (Dover & Murthi 2006: 5) This could be a potential retention measure, for example. Hence when a customer wants to unsubscribe, Customer Relationship Management might offer them to change their billing plan.

However, all the other costs listed in the quotation above might make rise the question about a consumer’s motivation to subscribe for a product - apart from the fact that a regular subscription might be cheaper than a pay-per-use model-, which is going to be covered by the next chapter.

2.2. The Motivational Background of Subscriptions

Though it might seem desirable at first sight, there is no universal customer segmentation. Many different studies show almost equally many different classifications, depending on industry, organisation or product. Indeed, “there are numerous ways in which online consumers [...] [can] be segmented.” (Conyette 2011: 93) Accordingly, it is neither feasible nor can it be the aim of this work to present a customer segmentation for subscription-based services. For this, the businesses referred to are too diverse. However, the following paragraph will highlight some aspects of what can generally motivate consumers to choose a subscription over a pay-per-use model.

Consumers spend time to find products to buy. Online, they spend even more time, as studies have shown. According to those studies, online customers spend more search time on all categories of products, both search goods and experience goods (cf. Huang, Lurie & Mitra 2009). This trend inspired researchers to coin the term of ‘informedness’ (cf. Clemons & Gao 2008). Up to now, customers devote a lot of time to finding exactly what they want (and are willing to pay premium prices for goods exactly suiting their needs), which resulted in the creation of the term ‘consumer informedness’. “This is not just consumers’ pursuit of products that are *better*, but rather *better for them*.” (cf. ibid.: 3) Given that, the possibilities of the internet favour the hunt for individually suitable products.

This is supported by Adjei et al., who write that “[i]ncreasingly, consumers use the internet as a vehicle for pre-purchase information gathering.” (Adjei, Noble & Noble 2010: 634) Hence, time is an important criterion in favour of consuming in the form of a subscription. Subscriptions save time. Time-consuming and laborious online search marathons can be avoided for the duration of the subscription, once having subscribed to an offer.

Other research articles approach a similar argument – beneficial effects - from a different angle.

Bhatnangar and Ghose, for example, distinguish between convenience shoppers and recreational shoppers (cf. Bhatnagar & Ghose 2004: 1353). Convenience shoppers are those who strongly consider criteria such as price and time spent, “availability of needed products, and convenience in parking and shopping.” (ibid.) All these criteria are covered by a subscription: the price is settled and known over a certain period of time, additionally a subscription might be cheaper than making single purchases, time has to be spent only once when the subscription is selected, the product needed is always available within the frame of the subscription, and convenience in parking and shopping are high because they do not occur. Though they do not speak of subscriptions, as it is not part of their study, the characteristics mentioned seem to be very well transferrable to an (online) customer of subscriptions.

Another insight into subscribers can be gained from the perspective of risk behaviour, also suggested by Bhatnagar and Ghose. They introduce a segmentation according to risk behaviour in internet shopping of eight different product categories. In this approach, they differentiate between product risk and security risk of internet shoppers, with the first referring to the risk that when making a purchase through an online store, the product cannot be physically examined; the latter meaning the risk of data leakage and related security concerns (cf. Bhatnagar & Ghose 2004: 1352). This is an interesting concept, which might be helpful to further illustrate the portrait of subscribers. In this case, one could say that product risk is minimised, because once the product is tested, the customer knows what they ordered. Additionally, though theoretically multiple product transfers occur, data transfer has to take place only at initiation of the subscription, which reduces security risk.

However, it cannot be concluded that a subscription customer is generally a risk averse consumer. A subscription usually binds for a certain period of time – during which product alternatives could occur that would better suit the customer's needs. Also, it is not common to make a first delivery or offer content access before a subscription has been signed, so product risk is still prevalent – it might actually be higher than in individual purchases, because when the subscribed service or product is not found the way the customer desired it, they are probably still going to be 'stuck' with it for the whole term.

The aspect of reliability and risk avoidance interestingly appears again in other studies of a subscription-related context. Chao describes it as follows: "Consumer subscription service entails transactions in two time periods: a forward transaction that allows consumers to select service options to hedge financial risks and a spot transaction that allows consumers to secure [the product or service] [...] on demand." (Chao 2012: 155)

This sums up consumers' main advantages. With subscribing, they have certainty with regards to price and access/availability, simultaneously they reduce search costs and enjoy the convenience of regular delivery or access.

After elaborating on the advantages on the consumer's side, the question might arise what is interesting about the concept from a provider's point of view.

Firstly, there are benefits such as the predictability of turnover, increased efficiency in production or economies of scale. Besides those arguments, subscriptions offer the unique opportunity of deep consumer insights, because they accumulate massive amounts of data on a regular basis and are able to link them to the demographic information stored on the individual customer's profile when signing up for the contract. Hence, they can perform data mining under superior conditions.

This introduces the next chapter on data mining, since this expression has become a hyped yet rarely explained phrase.

2.3. 'Swamps of Data'

In 2012, "more than 1.7 billion people, or almost one out of every four humans on the planet, [were] online." (Wasserman 2012: 13) This pervasion of information systems along with the 'the web 2.0' (cf. Hughes 2010: 418) allows companies to interact with their customers in unseen ways. It allows them to gather data from their websites and their apps, where people click, how long they stay on a page or listen to a song, they can register contacts with customer service centres, record usage times and catalogue age, gender, and financial details.

Meanwhile, the amount of data is excessive enough that "[e]nterprises are being swamped with data, and much of it is unstructured in origin." (Castellanos et al. 2012: 869) Furthermore, not all companies are proficient yet with interpreting this data.

Chordas, for example, wrote an essay about the use of data mining in the insurance industry and admitted that "there has been tremendous diversity in the results" (Chordas 2001: 117), meaning that some know how to process data – and some do not.

Nevertheless, those who have gotten a glimpse of what is possible with professional data mining, have realised that its possibilities are tremendous. Data mining can be the answer to a question formulated by Blank: " 'How many times have you thought about a theory or a hypothesis and said to yourself, "If I could just collect a million cases I would have the answer"?' " (Blank 2013: 463)

The answer to such a question could be, for example, which of my customers am I going to lose – and why? And if I knew – could I retain them? Because if known in advance, companies' customer relationship management could set off appropriate marketing measures, contacting customers on a personalised basis and consequently save precious profits. This has also been realised by research, for example by Ho and his colleagues, who formulated - though they actually referred to the publishing supply chain – that "e-commerce techniques [have the chance to] improve[...] customer relationship management." (Ho, Wang & Cheng 2011: 398) Stewart, Hess and Nelder also acknowledged that the possibilities of data mining to increase competitiveness have continuously been growing (cf. Stewart, Hess & Nelder 2011: 195-196). Still, those advantages do not yet define the term data mining itself.

In their 2011 work, Khajvand et al. provide a useful, though unfortunately very condensed definition:

> "data mining is the process of automatically discovering useful information in large data repositories. Data mining techniques are deployed to scour large database [sic] in order to find novel and useful pattern [sic] that might otherwise remain unknown" (Khajvand et al. 2011: 58).

This definition includes several aspects. Firstly, it is included that data mining has to extract information from large amounts of data. This is crucial. Secondly, though, the definition speak of 'patterns', which is a very broad expression. Thirdly, the definition mentions a data depository that is automatically searched for useful information, except that this data depository is usually called a data warehouse and composed of several parts. Since this explanation remains a bit vague due to the brevity in some aspects, the concept of data mining will be explained in more detail, starting with the building of a data warehouse. Still, it gave a good introduction to the topic.

2.4. Data Mining

Data mining has become the new Silicon Valley buzz word (cf. Geiselberger & Moorstedt 2013: 11) as well as one of the new trends of 21st century research.

But still, what exactly is data mining? This question will be explained in detail with the help of a recent introductory work by Rainardi, which was designed for practitioners wanting to build an own data warehouse in their company. Therefore, it is very well to grasp. Subsequently, the synthesis of subscription-businesses and data mining will then introduce churn management in section 3.

In order to define data mining, one has to start at the preceding point, which is the data warehouse. This “data warehouse is a system that retrieves and consolidates data periodically from the source systems into a dimensional or normalized data store. It usually keeps years of history and is queried for business intelligence or other analytical activities.” (Rainardi 2011: 1) As the citation explained, the data warehouse is the point of a company where all streams of data at an organisation’s disposal flow together.

Principally, this looks as follows:

Fig. 1: “Simplest Form of a Data Warehouse System”
Source: Rainardi 2011: 4.

The diagram above essentially depicts the most straightforward form of a data warehouse, because it has only one source system. Theoretically, an ETL can receive and bundle data from numerous source systems.

The source system “contain[s] the data you want to load into the data warehouse” (ibid.), which might be information on consumption behaviour of your customers, how they used the product, what they did on the homepage or in the app, or how many different numbers they called at what times of the day.

The second element is the ETL, the “extract, transform and load (ETL) system [which] then brings data from various source systems” (ibid.) together, in case there are several sources. Nonetheless, the data is not yet suitable for analysis when it is in the ETL system. Beforehand, it must be loaded into the target system. The mediating step via the ELT is necessary because it can transform and integrate data from different sources into one target system (cf. Rainardi 2011: 1).

This target system is the dimensional data store (DDS), which arranges data in a way that it is useful for analysis (cf. ibid.). In the case pictured above, the data warehouse actually consists only of the ETL and the DDS system, the source system is not part of it. (cf. ibid.: 4)

One more key feature of a data warehouse is the fact that it can store “very long history”, as opposed to transactional systems, which only store limited amounts of history (cf. ibid.: 10, 11).

Once the data is aggregated, it can serve multiple purposes, all of which are defined as Business Intelligence (BI) applications (cf. ibid.: 10, 12).

Business Intelligence in turn is defined as:

> “a collection of activities to understand business situations by performing various types of analysis on the company data [...] to help make strategic, tactical, and operational business decisions and take necessary actions for improving business performance.” (ibid.: 12)

According to Rainardi, BI activities are to be grouped into three main categories: Reporting, OLAP and Data Mining, as depicted by the figure below (cf. ibid.: 13).

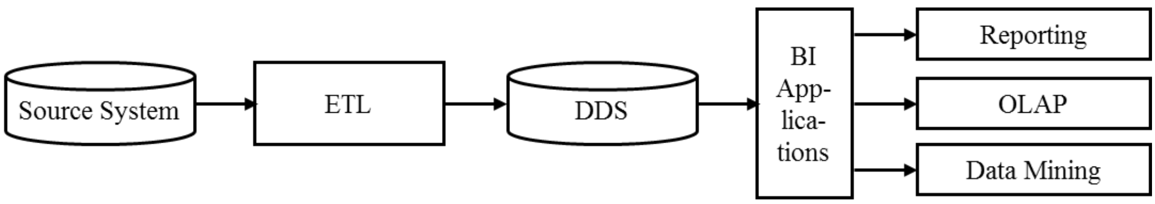

Fig. 2: The Purposes of Data from a Data Warehouse
Source: Own illustration, based on Rainardi 2011: 4, 10, 13.

Apparently, for this context the latter only is within the focus of interest. Reporting and OLAP are irrelevant for the topic of this work, which is why they will not further be explained. As shows the diagram, it becomes evident that data mining applications have to process very large amount of data, which indeed is one of its characteristics (cf. Khajvand et al. 2011: 58).

Given that, data mining is a BI activity and comprises the following activities: “data characterization, data discrimination, association analysis, classification, clustering, *prediction* [emphasis added], trend analysis, deviation analysis, and similarity analysis” (Rainardi 2011: 13). So far, the description sounds like almost every activity completed with data can count as data mining, which is why the term will first be defined in more detail. Secondly, it is also another reason why this paper concentrates on one main aspect, namely the prediction of behaviour. Additionally, behaviour prediction is the topic dominating research on subscription-based businesses and data mining, as will be shown later on.

In detail, Rainardi explains data mining as follows:

> “Data mining is a process to explore data to find the patterns and relationships that describe the data and to predict the unknown or future values of the data. The key value in data mining is the ability to understand why some things happened in the past and to predict what will happen in the future. When data mining is used to explain the current or past situation, it is called *descriptive analytics*. When data mining is used to predict the future, it is called *predictive analysis*.” (ibid.: 14)

This is a very precise and well understandable definition of data mining, which is why it was chosen. A central statement is that data mining makes use of historic data to predict what is likely to happen, which is exactly the purpose of churn prediction models. They analyse past customer patterns in order to give probabilities for customers on their leaving intention, differentiating those with a higher probability from others.

What is more, Rainardi not only names ‘predictive analysis’ as a concrete example for an application of data mining, but also customer profitability management (cf. Rainardi 2011). The topic of profitability is going to become of central importance for this work’s argumentation, and it therefore is going to be the second application of data mining this work will analyse. Meanwhile, data mining has namely found its use also in the research on Customer Lifetime Value-models, since the analysis of past buying behaviour allows for the prediction of future turnover to expect from a customer. The concept of

Customer Lifetime Value will allow to focus on the most lucrative subscribers, especially when acting under the condition of restraint marketing resources.

Now, as explained in the introduction, the crux in running a subscription based business is keeping the customer. Especially the industry of mobile telecommunication has a high interest in churn management, since they have to cope with cancellation rates of 20 – 40 percent (cf. Kim & Moon 2012: 11718). Additionally, it is more expensive to gain a new customer than to keep an existing one. This is why ‘prediction’ – the prediction of customers churn, to be exact – is currently taking very much space in data mining research. The following chapter will come back on the issue, defining the term ‘churn’ and outlining the reasons for creating churn prediction models.

However, as will be shown in subsequent chapters, churn management research has so far barely started to combine its findings with the application of CLV. Only recently, some studies have appeared that also advocate this approach. Yet even when combining those two methods, the studies still focus on comparing different predictive techniques, such as neuro-fuzzy techniques and neural network models (cf. Abbasimehr, Setak & Soroor 2013: 1279), instead of focusing on the development of strategic implications resulting from this combination. Those strategic recommendations can for example the adequate measures to be undertaken by Customer Relationship Management, in case high profitability and intention for termination coincide. In this case, the retention measures could even be elaborate and would still be economically effective, since they would secure maximum profit.

Therefore, it is the aim of this work to propose a framework for calculating the amount of profit that can be secured by applying data mining to subscription-based businesses. This will be achieved by combining churn management with the CLV-approach.

Beforehand, the third section presents not only a definition of churn management, but also a literature review on churn management research. This literary analysis mainly focuses on publications made since 2005, which is explained below.

This literature review was compiled not only to show the chances of data mining, but also to make proof of the circumstance that the studies forecasting customer switching behaviour are still very preoccupied with finding the best means of cancellation prediction, meanwhile maybe losing strategic focus. Subsequently, the second review will make proof of the research gaps that have opened up due to the thematic monopolisation on the telecommunications industry.

3. Churn Management

As outlined in the previous paragraph, churn management is the central topic of data mining applications for subscription-based businesses. Before progressing further, however, the term ‘churn’ has to be defined, because it is a term usually not encountered outside this discipline. Fortunately, researchers are quite concurrent on this term.

Hadden et al. describe it as follows: “[c]hurn management is the term that has been adopted to define customer turnover. More specifically, customer turnover is the concept of identifying these customers who are intending to move their custom to a competing service provider.” (Hadden et al. 2005: 2903) Hence ‘churn’ is a term describing the circumstance of customer’s cancelling their subscriptions.

This is expressed similarly by Lemmens and Croux, who define churn as “a marketing-related term that characterizes whether a current customer decides to take his or her business elsewhere (in the current context, to defect from one mobile service provider to another).” (Lemmens & Croux 2006: 276, 277)

When a customer decides to take their business elsewhere, it is crucial because of several reasons. Firstly, it is very disadvantageous to lose a (well-paying) customer to a competing business, because “[w]hen the number of customers belonging to a business reaches its peak, finding and securing new customers becomes increasingly difficult and costly.” (Hadden et al. 2005: 2903) At this point, Hadden et al.’s strategic recommendation is to give “higher priority to retain the most valuable, existing customers, than trying to win new ones." (ibid.) This is why "[c]ompanies need to understand their customers." (ibid.). In consequence, "[o]nce identified, the customers can be targeted with proactive marketing campaigns for retention efforts." (Hadden et al. 2005: 2903)

When speaking about churning customers, it has to be clarified that actually, there are different categories of leaving customers. This is subject of the next chapter.

3.1. Churning Customers

What is meant exactly when speaking of churning customers? Indeed, there are different groups of churners, according to circumstances and intentions. Hadden et al. give a very precise and thus useful definition of 'churners'. Following their categorisation, the diagram below illustrates the definitions and maps the progress of argumentation.

Since only subscribed customer can terminate a service, it has to be started with the existing customer base. Then follow those subscribers who decide to end the relationship with the respective company. They build a group of churners, or defection-prone customers.

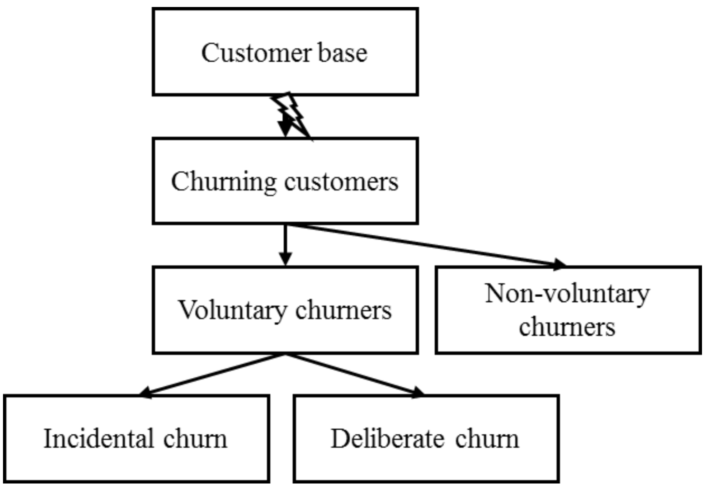

Fig. 3: Categorisation of Churning Customers
Source: Own illustration, according to Hadden et al. 2005: 2903.

According to Hadden et al. “[c]hurning customers can be divided into two main groups, namely voluntary and non-voluntary churners.” (cf. Hadden et al. 2005: 2903). Non-voluntary customers can easily identified, as they are those “who have their service withdrawn by the company [for example because of] [...] abuse of service and non-payment of service.” (cf. Hadden et al. 2005: 2903)

Voluntary cancellation in contrast is more complex, because it only occurs when a customer makes a conscious decision – which can be triggered by either external factors, such as a change of living circumstances, - or by factors that are of a more personal nature. Voluntary churners can therefore be divided into two groups, incidental churners and deliberate churners.

“Incidental churn happens when changes in circumstance prevent the customer from further requiring the provided service.” (ibid.) This can either be a change in financial resources, such as through losing employment, or a geographical change to a place where the service subscribed to is not available. However, “incidental churn usually only explains a small percentage of a company’s voluntary churn” (ibid.), which makes it on the one hand less interesting due to the small percentage. On the other hand, the customer *had* to terminate the service because they had unchangeable factors opposing its further use. In this case, it might not be unlikely that the customer would actually have continued the service, had not their external situation changed. As a consequence, the provider has neither reason nor chance to change or improve anything with the help of retention measures – except, e.g. to expand network coverage – nor would an adaptation in business model be successful. Consequently, non-voluntary and incidental churners are to be excluded.

Deliberate churn is more critical. In the case of a deliberate cancellation, a customer consciously “decides to move his/her custom to a competing company.” (ibid.) This means not only a reduction of the incoming cash flow, but it means also that the former customer’s streams of cash now probably flow into competing pockets – which is equivalent to a loss in market shares.

The decisive termination from customer side can have several reasons. First, “technology-based reasons, when a customer discovers that a competitor is offering the latest products while their existing supplier cannot provide them.

[Secondly] [e]conomical reasons include finding the product a better price from a competing company. [Other] [e]xamples [...] include quality factors such as poor coverage, or possibly bad experience with call centres, etc." (cf. Hadden et al. 2005: 2903) The last point is especially interesting, as it is something that could be discovered and prevented through adequate data mining and analysis – e.g., does (repeated) contact with customer support increase or decrease their probability to switch?

Of course, the two latter categories can overlap. For example, a customer has to terminate a service because of suddenly diminished financial resources, while still requiring some kind of this service. They then subscribe to a competitor that offers the same service at a better price (maybe at cost of quality, but that shall not be of central importance here). In this case, they would be an incidental churner first turning into a deliberately switching subscriber. Here it is crucial to note that of course economic reasons gain more weight in the change of a customer's financial circumstances - blurring the line drawn above.

However, besides this one transitional case, Hadden et al. provide an otherwise precise categorisation of churners. When referring to 'churners' from now on, it is clear it means the deliberate termination of a contractual relationship from customer side.

3.2. Churn Prediction Models: Relevant Factors

As the previous chapter illustrated, the reasons leading to the decision of decline are of great interest for determining the subsequent retention efforts. Therefore, it is worthwhile to investigate the variables that are of importance to create a predictive model.

This will be done in an exemplary way referring to a publication by Coussement and Van den Poel (2008). Their paper was chosen for several reasons. First, it is one of few that is held on newspaper subscription and secondly their paper is the only one giving a detailed list of all the explanatory variables included in the predictive model (cf. Coussement & Van den Poel 2008: 323). Thirdly, they are leading researchers on the topic, being two of very few who have made several publications on the issue.

Their variables derive from four different areas. First, the client/company-interaction, which gives, for example, information on the number of complaints or the time passed since the last complaint, but also on how the newspaper is delivered, and on the payment method. (cf. Coussement & Van den Poel 2008: 323)

Secondly, there are the variables related to the renewal of the subscription. For instance, "whether the previous subscription was renewed before" (ibid.) and "how many days before the expiry date, the previous subscription was renewed" (ibid.).

Thirdly, the so called socio-demographic variables are the variables describing the customer. They process information on age and gender, but also whether the subscriber is an individual person or a company. (cf. ibid.)

The fourth group encompasses variables such as monetary value, the length of the current subscription, what product the customer subscribed to, and during which month the contract is going to expire. (cf. ibid.: 323, 324) Those variables might become particularly interesting for this thesis because of the monetary value of a customer defined by their subscription. Consequently, this can be used to define how much effort and money is to be put into eventual retention measures.

For the interested reader, all the explanatory variables were made available in the appendix of their research paper (cf. ibid.: Appendix A, 323).

The process of determining the relevant variables is called feature selection. Hadden et al. (2005) provide a useful diagram to illustrate the process taking place when compiling a churn prediction model.

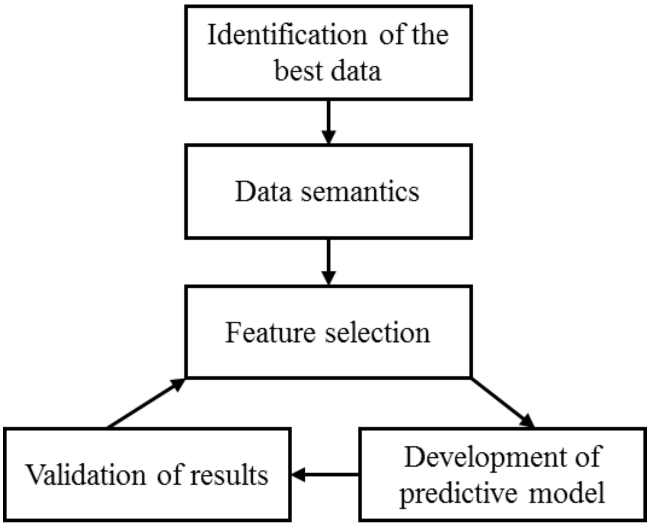

Fig. 4: “The Stages [of] a Churn Management Framework”
Source: Hadden et al. 2005: 2904.

The first step for data analysis needs to be data identification; because usually, as implies the term ‘data mining’, lots of data should be available, of which not everything is of use – provided that the data is saved and ‘historical data’ has been created (cf. Samimi & Aghaie 2011: 92). Therefore, “[i]t is necessary to identify the data that best suits the type of analysis that is being performed.” (Hadden et al. 2005: 2904) The seconds step is data semantics, meaning the understanding of what single pieces of data actually mean, because “data stored could be abbreviations of company specific terms, in numerical format, or having a dissimilar meaning to that which could be considered obvious.” (ibid.: 2905). In conclusion, “[d]ata semantics is the process of understanding the context of the data in a database.” (ibid.)

Thirdly, when it is clear what has been ‘mined’ so far, it is crucial to identify those pieces of data that best serve the upcoming analysis. This is called ‘feature selection’: it “is the process of identifying fields which are best for prediction [because] [...] it helps with both data cleansing and data reduction” (Hadden et al. 2005: 2905). According to Sun, Bebis and Miller (2004) it is a critical process and many research papers on data mining and churn prediction have shown lots of attention paid to choosing the right variables.

After this feature selection, Hadden et al. start building the model of defection prediction. Then they go into a circuit of validating the results. While not all studies have done this in their work and the models thus sometimes lack validation (cf. Bloemer et al. 2002), it is nevertheless an important point. The circle presents the chance to adapt the model, if necessary, through choosing new ‘features’ (data items) and adjusting it until the model manages the highest churn predictability.

This is a crucial gap separating theory from practice: for an organisation, knowing in advance which customers will terminate their service – when they will leave anyway - has no value in itself. Thus, it can be considered central for a churn prediction model to give the reasons for why customers want to terminate their subscription – something not all predictive models do (cf. Datta et al. 2001). Consequently, the organisation would have the chance to make an appropriate offer to retain the customer.

A first step in the direction of increased transparency was made by Coussement and Van den Poel, when they gave an overview of the most important churn predictors within their subscription-service setting based on the data of a Belgian newspaper publishing company (cf. Coussement & Van den Poel 2008: 314, 318). Interestingly, they said it might be important for marketing managers to know the important predictors for identifying likely churners (cf. ibid.: 314). “Consequently, it may be possible to adapt their marketing strategies based on this newly obtained information.” (ibid.). Nevertheless, giving marketing managers the criteria that influence the probability fir termination is not necessarily ensuring retention effectiveness, because correlation does not imply causation.

3.3. Predictive Accuracy and Interpretability

This is a point that was already taken up by Hadden et al.. They criticised that - at least in 2005 -, predictive models “predict customer churn, but [are] unable to provide an explanation as to *why* [emphasis added] a customer might churn.” (Hadden et al. 2005: 2909) Other predictive models, such as neural networks, are not helpful in that respect either, for their “classification rules are not output in an easily understandable form.” (ibid.: 2915) This was also claimed by other studies: “Neural network has good character in the predictive accuracy but lacks interpretability. This is very bad in the application fields, especially for the company that wants to know the aspects affecting the churn ratio.” (Liu, Qiao & Xu 2011: 1549)

So far, this is a weakness of predictive models. Either they do not to give an easily understandable reason at all, or the pieces of causal information they present are so complicated that it is highly difficult if not impossible for practitioners to understand and analyse them. Consequently, it becomes impossible to derive reasonable marketing actions. This is a gap that needs to be bridged. However, it may not have top priority, because once identified with a high probability to switch, a personal contact via phone call from CRM, for example, might easily find out what it is that dissatisfies the customer. Thus they can offer appropriate changes then.

Decision trees, in contrast, provide better interpretability, but according to Liu, Qiao and Xu, the output varies too much when the data is unstable: “Decision tree is a useful data mining technology for classification with good interpretability and sound predictive accuracy. However, it is unstable when the data varies, and this sensitivity makes prohibits it from being widely used.” (Liu, Qiao & Xu 2011: 1549)

This leads to the question: how do churn prediction models work?

Basically, they reduce the number of customers down to a limited number with higher probability not to expand their contract than others. This is usually referred to as ‘lift’: “Lift is probably the most commonly used prediction criterion in predictive modelling” (Neslin et al. 2006: 206)

In general, models use the term ‘top-decile-lift’, meaning that, for example in the case of Neslin et al., “the bestperforming models identify 10% of the customers who are three times more likely to churn than average” (Neslin et al. 2006: 206)

Leading researchers Coussement and Van den Poel make also use of top-decile-lift. In their article, it says the “top 10% decile is an evaluation measure that only focuses on the 10% cases most likely to churn” (Coussement & Van den Poel 2008: 317) and can be explained as follows: it is an "increase in density" (ibid.). This targeted and reduced number of customers then enables targeted marketing.

The following literary analysis, however, will show that churn research is currently focusing on methods instead of concrete application.

Concerning variable importance, the predictive models may oppose the criteria one might have set from a customer value point of view. According to the study by Coussement and Van den Poel, "monetary value and frequency [...] are not present within the top-10 list of most important churn predictors" (ibid.: 322). Nonetheless, monetary value might be a decisive criterion for practitioners when deciding on what to do in a particular customer's case.

In contrast, complaint management is highly influential, because the "recency of complaining [...] is also present in the top-10 most important churn predictors." (ibid.) This might be a useful point to bear in mind, as in previous research it was already stated that often, "companies do not deal successfully with service failures because most companies underestimate the impact of efficient complaint handling." (Coussement & Van den Poel 2008: 317) However, for the sake of brevity, the topic of complaint management will not be investigated any further.

4. Literature Review on Churn Prediction

In order to make proof of the topic that currently occupies churn research – techniques -, this chapter provides a first literature review of churn prediction. A second review presents the industries, which have provided the necessary data up to now.

With regards to literature review, determining its structure is crucial. Data mining in general is a topic that has only been surfacing since the turn of the millennium. This is valid in particular for the research on churn prediction models, which have only recently become subject of intense numbers of publications. This green light was also acknowledged by Liao, Chu and Hsiao, who published a literature review on data mining techniques, covering the decade from 2000 to 2011 (cf. Liao, Chu & Hsiao 2012). This choice of a starting point is supported by a statement of Keaveney and Parthasarathy, who wrote that in 2001, only very few studies, if any at all, had researched on churn prediction models (cf. Keaveney & Parthasarathy 2001: 375). Indeed, Keaveney and Parthasarathy did not yet present a complex prediction model in 2001, but mused on the general influence of customer loyalty. Moreover, they did not even use the meanwhile common term churn or a customer likely to churn, but spoke of “identify[ing] defection-prone customers” (cf. Keaveney & Parthasarathy 2001: 375). This shows that this topic was not an established field of research at the turn of the millennium.

Yet Hadden et al. published a literature review on churn management in 2005, which covered the “publications during the last five years” (Hadden et al. 2005: 2914). Their table also concentrated on the selection of the most precise prediction methods and therefore their paper included all relevant articles that had been published up to that point. Thus, the period from 2000 to 2005 is already well depicted. For this reason, this analysis focused predominantly on the period from 2005 onwards. The timeframe goes until 2013, in order to include the most recent publications.

The review is structured according to Liao, Chu and Hsiao (2012). As in Liao, Chu and Hsiao, key words and abstracts were used to identify relevant articles on churn management, all being published in significant academic journals (cf. Liao, Chu & Hsiao 2012: 11303). The articles were retrieved from the four most recognised and extensive online databases the university has access to: that is Science Direct, Springer Link, JStor and Sage Journals Online.

4.1. The Focus on Churn Prediction Techniques

Currently, many academic research papers on churn management do not focus so much on *what* can be found out, but rather *how.* This might be due to the fact that monitoring and predicting customer behaviour through data mining is a comparably recent BI Application.

Before presenting a summarising table, two examples of this concentration on techniques are briefly going to be explained. Samimi and Aghaie tried the suitability of ‘logistic regression formulation’ to analyse the usage rate of subscription-based businesses (cf. Samimi & Aghaie 2011). Coussement and Van den Poel tested “[a]n application of support vector machines while comparing two parameter-selection techniques” in order to predict customers’ termination of subscription services (cf. Coussement & Van den Poel 2008). Those two examples already give proof of the trend in data mining research to find appropriate measures.

It is a circumstance that has been acknowledged by Coussement and Van den Poel, who state that so far, all those “data-mining techniques [...] vary in terms of statistical technique (e.g., neural nets versus logistic regression), variable-selection method (e.g., theory versus stepwise selection), number of variables included in the model, time spent to build the final model, as well as in terms of allocating the time across different tasks in the modelling process” (Coussement & Van den Poel 2008: 313-314). This is still a proper description of the current situation, which is going to be underlined by the following table.

The review is structured as follows. To begin with, a similar analysis was conducted by Hadden et al. in 2005, which concentrated on the period from 2000 to 2005. They also compiled a detailed list of all works published on churn prediction (cf. Hadden et al. 2005: 2913) and sorted them according to statistical methods. This is why the table below mainly portrays the literary landscape during the period of 2005/6 to 2013. Two earlier examples are included due to reasons of consistency, because they are part of a latter analysis of data sets. The categorisation of techniques was completed according to Hadden et al..

Research article	Method						
	Cluster Analysis	Regression Analysis	Support Vector Machines	Neural Networks	Decision Trees	Social Network Analysis	Others
Abbasimehr, Setak & Soroor (2013)	X			X	X	X	
Ahn, Han & Lee (2006)		X					
Braun & Schweidel (2011)							Hierarchical competing risk model
Coussement & Van den Poel (2008)		X	X		X		
Coussement & Van den Poel (2009)		X	X		X		
Coussement, De Bock (2013)					X		Single algorithms, generalized additive models
Datta et al. (2001)				X	X		Genetic algorithms
De Bock & Van den Poel (2012)		X					Bagging, Random Subspace Method
Dover & Murthi (2006)							Stratified hazard model
Haenlein (2013)						X	
Huang et al. (2010)				X	X		
Huang, Kechadi & Buckley (2012)		X	X	X	X		Linear classifications, Naive Bayes, 'Evolutionary Data Mining Algorithm'
Hung, Yen & Wang (2006)				X	X		
Jahanzeb & Jabeen (2007)							*T* -test
Karahoca & Karahoca (2011)	X			X			x-means
Kim, Lee & Johnson (2013)		X					Partial least square optimisation
Kim, Park & Jeong (2004)							Factor analysis, reliability analysis,

							structural equation model
Lemmens & Croux (2006)							Bagging, stochastic gradient boosting
Liu, Qiao & Xu (2011)							Adjusted Real AdaBoost
Neslin et al. (2006)	X	X		X	X		Discriminant analysis, Bayes
Phadke et al. (2013)						X	
Samimi & Aghaie (2011)		X					
Verbeke et al. (2011)			X				Ant-Miner+ (basis: Ant Colony Optimisation)
Verbeke et al. (2012)		X					
Wong (2011)		X					
	Cluster Analysis	**Regression Analysis**	**Support Vector Machines**	**Neural Networks**	**Decision Trees**	**Social Network Analysis**	**Others**

Table 1: Research Articles and the Techniques Employed
Source: Own compilation, according to Hadden et al. (2005) and Liao, Chu & Hsiao (2012).

The table above shows that the discipline of defection prediction is still focusing on the task of testing and comparing different analytical methods. As can be seen, the variety of means employed is large. A new approach is the social network analysis, which was used in three different studies of five being published this year (cf. Abbasimehr, Setak & Soroor 2013; Haenlein 2013; Phadke et al. 2013).

However, this paper does not intend to investigate the techniques for model building in depth, nor to judge their suitability. Instead, it is the objective to present the chances of data mining for gaining consumer insights, and not how they are achieved by means of statistics.

Moreover, this question was already addressed by Neslin et al., who ran a tournament for predictive models, having given out a data set (cf. Neslin et al. 2005: 210). This way, they were able to compare many different approaches on the same data. Notwithstanding, they were still not able to identify the most useful model. On the contrary, they found that logistic regression and decision trees “perform relatively well”, but that it should be continued “to search for better techniques.” (Neslin et al. 2005: 209). In addition, Verbeke et al., as recently as last year, also concluded that the “issue of which classification technique to use for churn prediction remains an open research issue” (Verbeke et al. 2012: 212).

An even more recent work emphasises the findings outlined above (cf. Hashmi, Butt & Iqbal 2013). Unfortunately, Hashmi, Butt and Iqbal’s paper has not yet been published in a journal but only via ResearchGate, which is why an exact publication date is missing.[1] Therefore, the work is considered only briefly. It also presents a literature overview of the methods used in customer churn prediction including papers until June, 2013; but focuses only on telecommunications. In their study, decision trees are the most frequently used classification technique (cf. Hashmi, Butt & Iqbal 2013), but again, they do not write if it is also the most precise method. Despite the thematic limitation to telecommunications exclusively, it supports the relevance of this work and shows that the matter is up-to-date.

The issue with this meta-focus is that in the meantime, it is being neglected what can actually be achieved with this tool. Almost all of the conclusions of the papers listed above end with a final word on the accuracy of those means of analysis (usually including the necessity of more tests), whereas almost none gives an outlook on what these predictive models mean for concrete application in organisations.

Another over-concentration became evident when analysing the research collected above for the origins of the data sets. This is the issue of the next review.

[1] The publication seems to be forthcoming in the International Journal of Computer Science Issues, Volume 10, Issue 5, however, and went online at ResearchGate in October 2013.

4.2. Data Origins. The Focus on ‘the’ Churn Prediction Industry

As has probably already shimmered though, churn management research usually works with data from telecommunications, mostly mobile. This second table below was compiled to prove that churn research has a favourite industry: telecommunications. As can be seen, only 25% of the publications work with other than telecommunications data. The articles were researched as outlined earlier.

Research article on churn prediction	**Industry**			
	Telecommuni-cations	**Other Subscription-based businesses**	**Other**	
Abbasimehr, Setak, Soroor, (2013)	X			
Ahn, Han & Lee (2006)	X			
Coussement & Van den Poel (2008)		X		Newspaper subscription
Coussement & Van den Poel (2009)		X		Newspaper subscription
Coussement & De Bock (2013)			X	Online gambling industry
Datta et al. (2001)	X			
De Bock & Van den Poel (2012)			X	European bank
Dover & Murthi (2006)			X	Online content provider
Haenlein (2013)	X			
Huang et al. (2010)	X			
Huang, Kechadi & Buckley (2012)	X			
Hung, Yen & Wang (2006)	X			
Jahanzeb & Jabeen (2007)	X			
Karahoca & Karahoca (2011)	X			
Keaveney &			X	Online services (e.g., America

Parthasarathy (2001)				Online, Delphi, Prodigy, MSN, CompuServe)
Kim & Moon (2012)	X			
Kim, Lee & Johnson (2013)	X			
Kim, Park & Jeong (2004)	X			
Lemmens & Croux (2006)	X			
Liu, Qiao & Xu (2011)	X			
Phadke et al. (2013)	X			
Verbeke et al. (2011)	X			
Verbeke et al. (2012)	X			
Wong (2011)	X			

Table 2: Research Articles and the Data Origin's Industry
Source: Own compilation.

In turn, this opens up immense research opportunities to apply churn prediction models to other industries, of course. A first step in that direction is a research published in 2013, which applied churn management to the online gambling industry (unfortunately). They worked with real life data from online poker players at 'bwin' (cf. Coussement & De Bock 2013). Apart from the moral dubiousness of this project, all the online subscription-based businesses for media content named at the beginning are probably at least equally interesting opportunities for mining consumer data and for testing the predictive concepts developed previously.

This research suggestion introduces the next chapter, which summarises and discusses the findings made so far and outlines research opportunities in more detail.

4.3. Research Gaps and the Development of a Framework

Thus far, the following milestones have been achieved: a first research gap was unravelled and the foundations for a new concept have been laid that will allow quantify the beneficial effects of churn management and CLV-selection.

The research gap is related to the situation described above. So far, cancellation studies worked with data from telecom operators in a huge majority of cases. Except for the study on online gambling, it has not recently been applied to digital businesses. The one study on online content providers might seem like it, but it was already conducted 12 years ago, speaks of the internet revolution as a current topic and mostly uses the term 'customer switching behaviour'. (cf. Keaveney and Parthasarathy 2001) Still, it was included in the table above given that it is one of the first research papers that investigates the chances of developing a churn prediction model, and that it used different data than mobile communications data. However, it cannot be considered a study at the same level as the studies and predictive models that have been developed over the last five to seven years.

Therefore, this work proposes to apply recently developed models to one of those subscription-based online media content providers described in chapter 2.1., for example. Depending on the maturity of the business, they might run broad and useful data mining. This 'interdisciplinary' application of churn prediction might bear interesting new insights and advance the whole field of studies.

Secondly, it was spoken of a framework. This idea shall now be explained. Up to now, a considerate number of scientific papers on churn has been published, which could be seen in the literature reviews of the two previous chapters. However, as was already argued, it is no advantage in itself for a company to know in advance, which customers are going to terminate their subscription, when nothing result from this knowledge. This is a crucial point.

Basically no churn researcher has taken their model to the next level, where it becomes really interesting for organisations, namely the consequential actions that are to be derived from these predictions. Usually, the models developed take the way described by Hadden et al. (2005) and which was displayed in figure 4: the data items are chosen, the model developed, then it is readjusted, maybe data is in- or excluded, the model is tested again, and then research is finished (please cf. figure 4, p. 29).

Therefore, it is going to be explained how companies can progress after the determination. The aim is clear: to retain the customers at risk. This is the task of Customer Relationship Management (CRM). Because the department knows the customers, they can run targeted and direct marketing measures. Yet the resources of the department are usually limited. Hence they will not be able to address every single customer. Moreover, not every customer is worth retaining. Therefore, they ought to be selected; ideally according to their value, so the retention measures can be targeted at those customers that bear the highest profits. This selection or segmentation can be done using the Customer Lifetime Value concept. This model is meanwhile also built by using big data, so it is another new and interesting chance of data mining for gaining consumer insights. Nevertheless, this concept has so far not been applied to subscription-based businesses, which will be shown in detail with the means of a literature analysis in the following section.

This introduces the next and final section, which describes the framework outline above in detail, explains the Customer Lifetime Value concept along with a second literature analysis and shows how an organisation has to progress with the customers at risk. Therefore, the next section will also briefly outline the importance of retention measures and the role of direct marketing.

5. A Framework for Quantifying the Effects of Data Mining

So far, no research in the context of predicting customer defection has presented a way to quantify the advantages of data mining. Only one paper calculated the monetary difference that is affected by increasing the predictive accuracy of a churn model (cf. Neslin et al. 2006), but this paper focused on the delta between well and better performing models, not on the all-over benefit. Thus, this chapter will now set a basis that might motivate companies to invest (more heavily) in data mining. Returning to the framework described previously, the process following churn model development would then look as follows, based on the chart by Hadden et al. (2005):

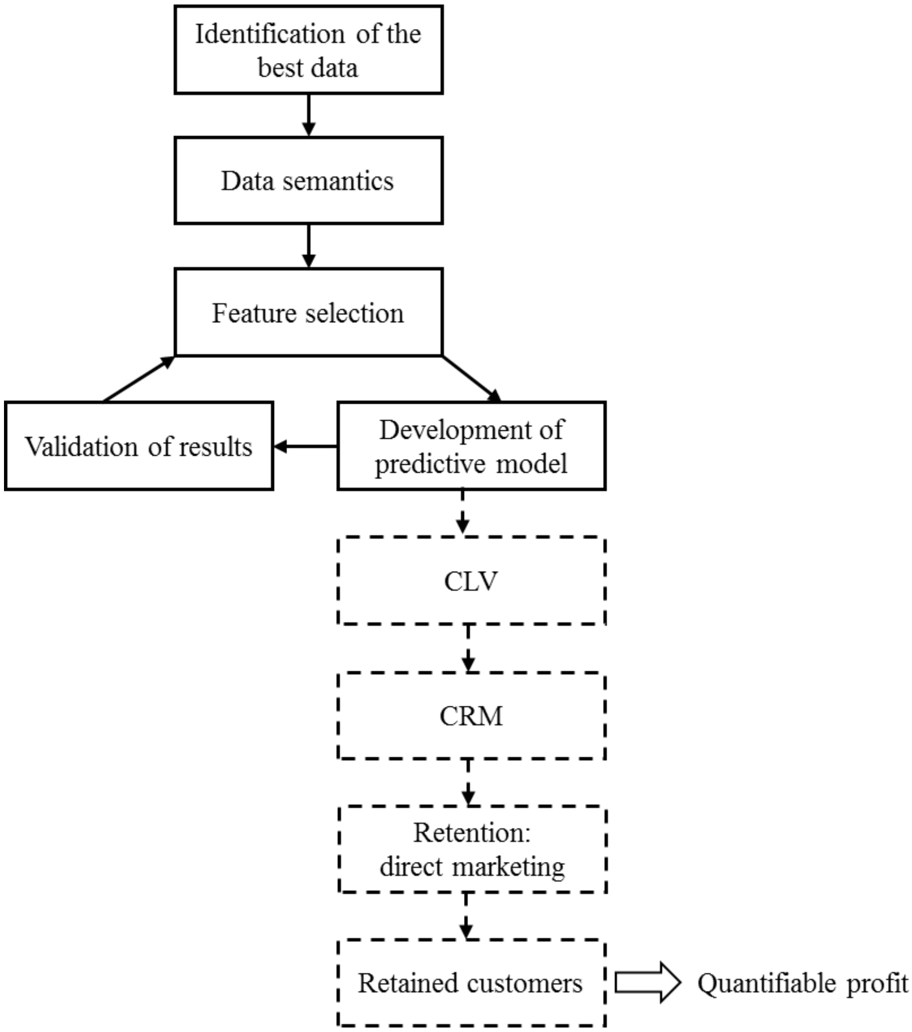

Fig. 5: The Stages of a 'Return on Data Mining'-Framework
Source: Own development, based on Hadden et al. 2005: 2904.

The model above is structured as described before. After the predictive model defined the customers likely to leave, the number has to be reduced. This follows from the fact that "customer retention efforts have also been costing organisations large amounts of resource." (Hadden et al. 2005: 2902) However, "through prediction, they can focus on a small scale of customers, and pool resources to conduct customer retention with less costs." (Liu, Qiao & Xu 2011: 1549) Yet this reduction can even be targeted further, because "not all loyal customers are profitable and not all profitable customers are loyal” (Kim & Moon 2012: 1179), which means that not all customers are worth retaining.

This further targeting is then done selecting those with high customer value, which can be determined with the help of a CLV-model. This group of customers with low staying intentions and high margin is then handed over to the CRM department in order to run retention marketing, convincing them to prolong their subscription. As soon as the customers of a certain period are retained (or not), their respective new contract’s margins’ net present value can be calculated. This number then quantifies what has so far been referred to as ‘return on data mining’. It is the amount of profit that was secured for the sake of the organisation because the customers, who would have left otherwise, could be convinced to stay.

Often, the implementation of data mining-models faces various thresholds and practical constraints like budget limitations (cf. Prinzie & Van den Poel 2005: 630). This framework, which renders the results of data mining calculable, might also help companies expand or justify existing data mining applications. In addition, it might even support other companies that have not yet started with data mining, due to various thresholds, to overcome their hurdles.

Moreover, customer retention was done highly efficiently, because the efforts were only focused on lucrative candidates. This leads back to the concept of CLV, which is subject of the following chapter.

5.1. Customer Lifetime Value

Meanwhile, with the spreading of data mining applications, Customer Lifetime Value has become one of *the* segmentation techniques in marketing. The trend has even gone so far that homogeneity is not a top-important criterion for customer segmentation anymore, rather it is customer value (cf. Chan 2008: 2754, 2755). This is due to the many advantages of a segmentation using customer value:

> “Such questions as what sort of marketing strategies should be preferred for which customers, how much investments should be made for them and which marketing campaigns should be followed can all be determined by calculating lifetime value of customers.” (Hiziroglu & Sengul 2012: 766)

This framework employs this concept in order to determine if marketing measures should be undertaken at all. Afterwards, it can also be used as a guideline to limit the height of the ‘investment’, meaning what costs for retention measure are appropriate.

The idea of assessing every customer with a certain value has spread enormously in research of the last decade. Consequently, there is a wide range of ways to define a customer’s ‘lifetime value’. Nevertheless, this chapter will sort out the most useful definition.

The number of CLV-models has risen (cf. e.g. Hwang, Jung & Suh 2004), so that in 2012, Hiziroglu and Sengul presented an extensive summary of all different kinds of CLV-models (cf. figure below).

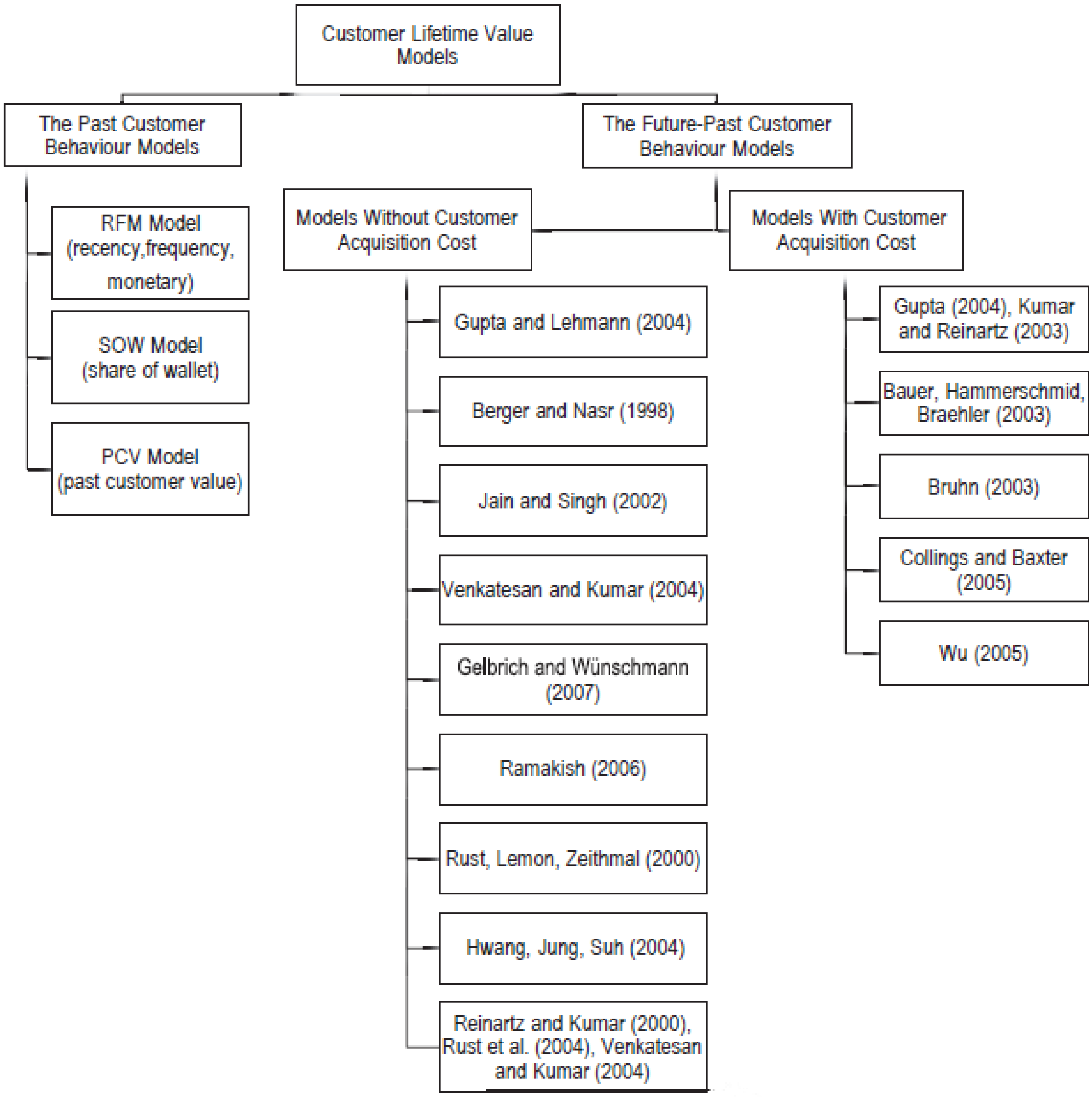

Fig. 6: "Classification of CLV Models"
Source: Hiziroglu & Sengul 2012: 767.

The chart above provides a summary and structures CLV-models in two main categories: those based exclusively on past contributions and those based on both past and future returns. Since the model is oriented towards future retention, the second category is the more suitable one. This is supported by the fact that "the results indicated that the model that represented the future-past customer behaviour model class as found to be superior" (Hiziroglu & Sengul 2012: 766)

As can be seen, the category of future-past returns is subsequently split in two sub-categories, one including acquisition costs and another one excluding them. With this CLV-approach finding its application before the retention measures are decided upon, the model has to exclude acquisition costs. (cf. Hiziroglu & Sengul 2012: 767) Additionally, as the CLV shall also serve as a defining value for the height of retention costs, they can of course not be included in the model.

Within this range, the definition by Gupta and Lehmann based on their work "Customers as Assets" is considered to fulfil the premises of this work (cf. Gupta & Lehmann 2003).

They define that "Customer lifetime value (CLV) is the present value of all future profits generated from a customer." (Gupta & Lehmann 2003: 10). Referring to the point they make - that it is common practice to estimate the time a customer will stay with the company – acting in the field of subscriptions should be advantageous here, for subscription contracts often have a maturity, except they are on a monthly basis. But even in that case, companies have probably quite useful statistics on how long customers of a certain subscription type stay on average, regardless of a maturity date.

The respective equation looks as follows:

$$\boldsymbol{CLV} = \sum_{\boldsymbol{t}=\boldsymbol{1}}^{\boldsymbol{n}} \frac{\boldsymbol{m_t}}{(\,\boldsymbol{1}+\boldsymbol{i})^{\boldsymbol{t}}}$$

m being the margin the customer is going to contribute, *t* the time they are going to stay and *i* the discount rate (cf. ibid.).

To reinforce this choice, it must be said that many other works define CLV similarly, at least with respect to using only the margin. For instance, Bolton, Lemon and Verhoef also define CLV as "as the net present value of *all* earnings (i.e., revenues less costs) from an individual customer (e.g., Berger and Nasr 1998; Dwyer 1989; Gupta, Lehmann, and Stuart 2001; Rust, Zeithaml, and Lemon 2000)." (Bolton, Lemon & Verhoef 2004: 271-272). As can be seen, they also already sorted through a vast array of literature, coming to the conclusion that it is the net present value of all predicted earnings (cf. also Ngai, Xiu & Chau 2009: 2595) after the deduction of all costs (cf. Kim et al. 2006: 102), so profit only.

With regards to the framework, customers at risk were now sorted according to their future profit contribution. Then, “[c]orrect calculation of CLV can facilitate a firm classify its customers based on their lifetime value rankings so that different marketing strategies can be developed for each group” (Hiziroglu & Sengul 2012: 766). This is the task of Customer Relationship Management, though in this case, there is only one - heterogeneous - group of customers, and so it should rather say ‘to develop a targeted marketing strategy for each customer’. Nevertheless, this introduces the next chapter, which is on CRM.

Beforehand, a literature analysis will provide the chance to see the thematic focus of CLV research, similarly to the analysis held earlier on churn data. This will show that lillte research has been done so far on determining CLV in the context of other subscription-based businesses except for a small number of telecommunication studies.

5.2. Literature Review on the Concept of Customer Lifetime Value

As it was interesting to see the thematic focus of cancellation management research, it is equally interesting to realise the thematic focus of CLV research.

According to the reasons explained in detail in chapter 4, this review covers the last decade, starting in 2003. The table is structured as the summary on the data origins in churn research, to achieve better comparability.

Research article on Customer Lifetime Value	**Industry**			
	Tele-communica-tions	**Other subscription-based business**	**Other**	
Benoit & Van den Poel (2009)		X		Financial services (“contractual setting”)
Chan (2008)			X	Nissan automobile retailer
Chan, Ip & Cho (2010)			X	Company that offers professional electrical products to construction industry and individual households

Glady, Baesens & Croux (2009)			X	Belgian retail financial service company
Glady, Baesens & Croux (2009a)				Retail banking sector
Gupta & Lehmann (2003)			X	Ameritrade (American online broker)
Haenlein, Kaplan & Beeser (2007)			X	Retail banking
Han, Lu & Leung (2012)	X			Telecom services in China
Hiziroglu & Sengul (2012)			X	Firm which engages in catalogue marketing in the US
Hwang, Jung, Suh (2004)	X			wireless communication company in Korea
Khajvand et al. (2011)			X	A health and beauty company that manufactures shampoos, soaps etc.
Kim et al. (2006)	X			wireless telecommunication company

Table 3: Research Articles on CLV and the Data Origin's Industry
Source: Own compilation.

Chan (2008) also did an overview of then recently conducted studies on customer segmentation (cf. Chan 2008: 2755; Table 1, "Recent customer segmentation research summary"), which features the same number of studies. His results show only two examples of segmentations in subscription-businesses, which are, unsurprisingly, on mobile communications. (cf. ibid.) This supports the findings of the table above. One of them is listed in the table, anyway. The first study listed is on the car insurance business, which might be considered a subscription-based business and is therefore categorised as such. However, except for one debatable example, the subscription-based industry apart from mobile telecommunications is left blank. This, again, bears promising research opportunities.

After this venture into the field of global research, it shall now be returned to the framework outlining what to do with the customers predicted to end their contract. As the customers with high CLV were selected, the next step is to be executed by Customer Relationship Management. Therefore, the next chapter is on CRM.

5.3. Customer Relationship Management

Meanwhile, companies have started to "shift away from their traditional, mass marketing strategies, in favor of targeted marketing actions." (Coussement & Van den Poel 2008: 313) This is where Customer Relationship Management (CRM) comes into focus. An important task of CRM is the "proper understanding of customer behavioural patterns" (Samimi & Aghaie 2011: 90). In order to understand the respective customers, it has been common to segment customers into different groups in order to build typologies that help imagine a certain 'kind' of customer and thus support marketing functions in developing adequate retention strategies, "product differentiation strategies or marketing mix allocation strategies" (Samimi 2011: 90). "To obtain such valuable knowledge [on customer behaviour], organizations use segmentation techniques to differentiate customers based on their usage patterns." (Samimi 2011: 90) One the one hand, this was done using the CLV approach, but it resulted in a theoretically very heterogeneous group. On the other hand, other segmentation techniques might become superfluous with the help of churn prediction models giving reasons for why a customer is intent on leaving the company.

More generally speaking, "CRM, a recent marketing paradigm, pursues long-term relationship with profitable customers." (Kim et al. 2006: 101) This is an approach also conveyed by Rainardi: CRM "is the activity of establishing contact and managing communications with customers, analysing information about customers, campaigning to attract new customers, performing business transactions with customers, servicing customers, and providing support to customers." (Rainardi 2011: 441) As it reads, the tasks of CRM are numerous.

Therefore, the tasks and purposes of CRM are illustrated by the following figure:

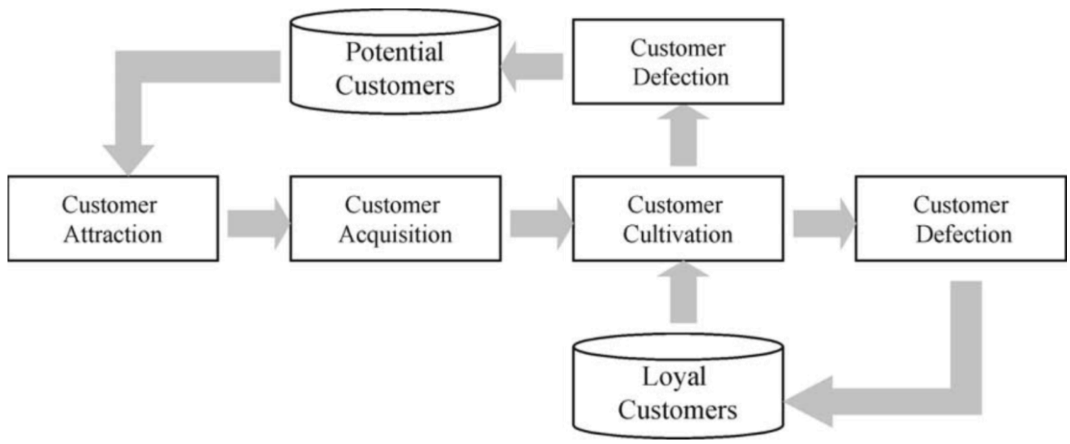

Fig. 7: "The Scope of CRM"
Source: Hwang, Jung & Suh 2004: 182.

As in the quotation by Rainardi, the figure shows that CRM has to serve a considerable amount of purposes – attracting new customers, taking care of the existing ones and hindering those intent on terminating their contracts to actually leave. It can also be seen that in 2004 customer churn is not yet called as such but runs under the tag of 'customer defection'.

Moreover, in Rainardi's approach, CRM processes lots of data on their customers, including analyses of personal and demographic information as well as the prediction of customer behaviour (cf. Rainardi 2011: 441). Yet, resources are restrained; this leads to the need to focus. One way is to concentrate only on those customers who are not planning on extending their contract – as predicted by a model - and who, ideally, provide the company with a considerable margin through their subscription. This is the reason for the order of tasks in the framework.

This is supposed to help CRM focus on fewer but more effective tasks – those tasks being the retention measures to retain the 'defection-prone customers', because, as explained by another study, "it is possible that many of those customers [predicted likely to churn] renew their contracts if appropriate retention promotions are offered at the right time through the right marketing channel." (Kim & Moon 2012: 11718)

This quotation shows that retention measures must be purposefully targeted and cannot be conducted in a mass marketing form. Instead, the promotions must be adapted and have to come through the right channel at the right time. This is something only direct marketing can achieve, as already indicated by references to Coussement and Van den Poel at the beginning of this chapter. Some studies even recommend organisational adaptations: "Telenor should specifically focus on customer retention and not just on acquisition phase. It should establish a separate retention / churn management department to deal with this emergent issue in Pakistan's telecommunication industry." (Jahanzeb & Jabeen 2007: 129). Again, the task of this retention/churn department would be the execution of direct marketing, which is why the organisational details are not of interest in this case.

The next chapter two chapters will present the concept of direct marketing and retention.

5.4. Direct Marketing

In research as well as probably in practice two kinds of marketing are distinguished: direct marketing and mass marketing (cf. Bose & Chen 2009). Mass marketing is possible without having made any detailed data analyses, for it is strewn out to the masses, usually via "television, radio, magazines, and newspapers." (Bose & Chen 2009: 1) It is defined as follows: "Mass marketing targets large groups of customers. It does not discriminate between customers within a group and the information delivered to customers is uniform." (ibid.)

Hence, the content that is spread through uniform channels is not individually adapted to different consumer's needs, contrary to what data mining and prediction models enable to.

In contrast to that, direct marketing is "different from mass marketing in that it targets individuals or households. Different customers are subjected to different marketing information." According to Bose and Chen, direct marketing is defined as: "communications where data are used systematically to achieve quantifiable marketing objectives and where direct contact is made, or invited, between a company and its customers and prospective customers" (ibid.).

This is exactly what this work's framework suggests: the use of data mining to find out about every single customer's intention, to select those who plan to leave, subsequently selecting those with considerable CLV so the efforts remain economic. On this number of customers, which was reduced twice, CRM can then launch targeted marketing measures.

For those direct contacts, it might be interesting to know in which ways customers behave during their subscription, in order to prepare adequate retention offers.

Unfortunately, research on customer's behaviour during a subscription is rather limited. However, some papers give hints on what could be found out on customer behaviour during a subscription. "Different groups of customers possess different degrees of intention to purchase. Therefore, a probability model can be used to predict the propensity of repurchase, given the class of a customer." (Samimi & Aghaie 2011: 91) This could, for example, result in insights into what customers might do with their subscription. For example, in case of a mobile phone provider, this could indicate the likelihood with which a customer is willing to subscribe to a second contract, for example for a family member.

Another alternative topic might be the behaviour in case of staggered subscription plans. Those are often applied in the media industry, where criteria such as advertisement versus ad-free, breadth of media accessible and number of devices that the service can be used on can vary according to the subscription plan. The propensity to repurchase could in this case reflect the likelihood with which an existing customer is willing to upgrade their subscription from a free and ad-supported service to an ad-free service, or from a medium programme to an advanced programme with access to more devices and offline offerings. This 'increased interest' might also be what Samimi and Aghaie refer to later on, when they state that their model "can be [...] employed [...] to recognize particular groups of customers whose tendency towards service usage has changed (either in a decreasing or in an increasing direction)" (Samimi & Aghaie 2011: 97).

One paper on churn in the telecommunications industry gives brief insights into what the "most common retention programs among telecommunication service providers" (Kim, Lee & Johnson 2013: 2198) are. One measure to convince customers is "a financial incentive that allows them to purchase a new mobile device at a deeply discounted price." (ibid.) The second option described is basically a more-friendly-than-usual call centre service: "Another popular retention program is to simply provide a better customer call center service in regards to billing and call quality peacefully through educated and experienced receptionists to resolve many questions and complaints and enhance customer satisfaction and loyalty" (ibid.: 2198, 2199). The only thought given on those two possibilities is that an offer to sell a new phone at a greatly reduced price will probably cost more than trying it with a better call centre service (cf. ibid.: 2199).

With regards to the costs mentioned, there is another financial factor that can be considered by a customer when thinking about cancellation: switching costs. Indeed, it might cost a customer considerably to switch to a new company. As explained in the beginning, a subscription is usually composed out of several price factors, one of them being for example an activation fee, which will probably occur again when the consumption is shifted to another provider. The concept of switching costs is explained in the next chapter.

5.5. Switching Costs

Costs were already mentioned in the context of subscription pricing. Time spent on searching the product desired can motivate a consumer to make a subscription in order to reduce these search costs. Similarly, high switching costs can keep a subscriber from terminating the service.

Therefore, this chapter will give a short overview of the concept of switching costs and other factors that might refrain subscribers from changing.

Switching costs are “defined as the cost occurred when a customer replaces the current supplier with another” (Min & Wan 2009: 109). These costs can occur because the customer needs to search for a new provider, or because the subscriptions overlap a certain time, leading to double payments. Switching costs can also be a new activation fee when starting a subscription with a new provider and “can deter customers from using these suppliers” (Lam et al. 2004: 294). However, true to their nature, they cannot be influenced by the provider that is still in charge of the subscription, which is why switching costs have to be excluded as a retention measure.

Another interesting aspect is the circumstance that switching costs are an argument that opposes the usage of loyalty as indicator for churn behaviour: “even satisfied and loyal customers may leave the current service provider if they find attractive goods and services at more affordable terms from other service providers.” (Kim & Moon 2012: 11720) Customers might also leave because loyal customers can be multi-loyal to several brands, or because they are only ‘loyal’ due to convenience or high, otherwise occurring, switching costs (cf. ibid.).

Cancellation fees are another form of switching costs. According to Fruchter and Sigué, these fees are used by T-Mobile and AT&T, but also in the banking sector. Moreover, the companies report positive effects.

> “T-Mobile, and AT&T claim that, thanks to cancellation fees, they have been able to limit customer attrition and provide phones at lower costs to customers who enroll in some specific long-term plan (Kharif, 2008). Cancellation fees are also common in the financial industry. For example, in addition to subscription (membership) fees, Visa now charges some of its member banks cancellation fees to prevent them from leaving” (ibid.: 2181).

Activation and cancellation fees are ways to retain customers so they do not decide to cancel. However, the framework targets customers who have the intention not to prolong their contractual relationship. Additionally, those fees cannot be imposed on a customer after the signing of the subscription contract. Hence, switching costs cannot be part of retention programmes, except for the fact that CRM can offer the annulment of cancellation fees for the extended contract. In the case of an extension, activation fees usually do not occur.

5.6. The Descriptive Framework

The steps following the setup of a churn predictive model have been explained. In this chapter, they shall be assembled in order to outline the output and the descriptive nature of the framework.

Subsequent to the forecasting of switching intentions, which gives a certain number ($\boldsymbol{n}$) of customers at risk, this output is filtered by the application of a CLV-model. This step reduces the number of defection-prone customers down to those worth retaining ($\boldsymbol{n_{CLV}}$). This data-set of customers is then handed on to Customer Relationship Management, who take care of it by starting retention measures. Ideally, they would be able to retain all, which is highly unlikely. It is assumed that the customers predicted to churn are all correctly predicted, and that no retention effort is 'wasted' on customers who actually did not want to cancel their subscription.

In the final step, the number of those who extended their contracts is assessed. Additionally, the cost for each retention measure must be calculated. In this procedure, the margin of the respective contracts can also be identified. These margins can then be taken to calculate the net present value of the profit that was secured. Subtracting the costs for retention and adding up all profits at net present value, the number gives the quantifiable success of data mining for an organisation. This is a new approach, which has in not yet been suggested by research and which can be considered key to this master's thesis.

Organisations that have been hesitant so far about how to proceed with data mining in their own domain might be convinced to expand and invest in data warehousing, data mining, churn prediction and CRM, because the benefits of these expenditures can be quantified by a number.

A suggestion for future research, in this context, is to develop a predictive model based on this descriptive one. It might be interesting to know in advance which amounts are at stake. For the purpose of creating a model capable of estimations, retention probability is the decisive factor to be determined.

However, retention probability is a topic very few research findings exist on, as the next chapter shows.

5.7. Retention Probability

Retention probability is the variable missing to build an estimative model out of the framework developed so far. Retention probability describes the all-over probability of changing the mind of cancellation-intent customers and the effectiveness of single retention measures.

The advantage of a high retention rate lies within the positive effects on profitability: "The consequences of enhanced customer loyalty in service firms are increased revenue, reduced customer acquisition costs, and lower costs of serving repeat purchasers, leading to greater profitability" (Lam et al. 2004: 293).

According to studies quoted in Kim & Moon, for example, this effect on profitability can be calculated. They claim that "companies can increase the average net present value of a customer by at least 35% or even up to 95% by merely boosting the customer retention rates by 5%. [Another study] also concluded that a 1% improvement in retention can improve customer profitability by about between 5 and 50 times than a similar improvement in margin and acquisition cost." (Kim & Moon 2012: 11718). Other researchers argue similarly: "For example, a 5% increase in customer retention in the insurance industry typically translates into 18% savings in operating costs" (Bhattacherjee 2001: 352)

Coussement and Van den Poel summed it up quite appropriately, stating that "[i]t has been shown that a small change in retention rate can result in significant changes in contribution" (Coussement & Van den Poel 2008: 313).

However, it has to be noted that none of those papers gives concrete numbers for retention rates. They speak of the effects of an increase, but concerning the retention itself, nothing is given away.

A paragraph on the possible (monetary) gains inherent to churn forecasting could only be found in Neslin et al. (2006). They explain that churn prediction enables to know which customers are at risk so they can be addressed, and that besides, one also gets to know which are not – consequently, the company does not waste money on customers who were not intent on leaving anyway (cf. Neslin et al. 2006: 204).

In addition to those advantages, they suggest that the more accurately a churn model is able to estimate, the more money can be saved by the subsequent retention campaign. Then they try to quantify this statement. Increased predictive accuracy could “easily amount to changes in profit in the hundreds of thousands of dollars by using one method rather than another.” (Neslin et al. 2006: 208). To show that, they make an exemplary calculation, which remains only hypothetical for several reasons. First, they do not focus on the most valuable customers, but on customers all across the CLV-range, calculating with a range of values: “$500, $1,500, and $2,500” (cf. ibid.: 207).

The crucial point, retention probability, is only referred to in a single sentence: “No published data are available for acceptance rate (γ), so we investigated 10%, 30%, and 50% rates.” (ibid.) The definition for the acceptance rate (γ) is: “γ = the fraction of targeted would-be churners who decide to remain because of the incentive (i.e., the success rate of the incentive)” (ibid.: 205). The key statement, however, is that no published data is available on the topic of retention rates.

This lack is probably due to the circumstance that retention rates are difficult to estimate. Gupta and Lehmann explain that “[p]ractically, retention rate is one of the most difficult metrics to empirically estimate.” (Gupta & Lehmann 2003: 11) This difficulty is usually, as in Neslin et al., bridged by assumptions: “Therefore many applications either assume a retention rate or estimate a retention rate that is constant over time” (Gupta & Lehmann 2003: 11-12). The latter assumption is not realistic because the rate might change over time and the longer a customer stays with a company, the higher might become their retention probability due to high loyalty.

An interesting point made by Kim and Moon is the circumstance that estimative models seem to assume that “all the customers will identically respond to a retention promotion or customers with higher churn probability are more likely to accept retention promotion offers.” This shows great need for empirical investigation, because it might rather be expected that the greater the probability to churn, the greater the obstacles to retaining the customer. Indeed, the "probability of accepting retention marketing promotions may be dependent on several factors such as switching costs [...], customer satisfaction [...] and several demographic variables" (Kim & Moon 2012: 11719). But as Kim and Moon argue, the probability *may* depend – but it is not proven yet. This was also realised by Chan, who claims that "[u]nfortunately, the link between customer segmentation and marketing campaign is missing." (Chan 2008: 2754)

To sum it up, no study has been published on retention rates so far. Neslin et al. closed the research gap by using estimations. Nevertheless, in order to be able to reliably calculate the gains of a retention campaign, it is necessary to develop a formula that measures the probability of success of certain marketing measures on certain customers.

Of course, this rate will be different from industry to industry, and probably also from company to company, but maybe similarities will occur according to means of retention and groups of customers. Additionally, it will enable future researchers to quantify 'returns on data mining' in advance.

6. Research Suggestion: A Predictive Framework

As the previous chapter has outlined, a suggestion for future research is to take the descriptive framework for measuring the gains of data mining and turn it into a predictive model.

Based on the framework introduced in chapter 5, and integrating the retention probability described, the model would look as follows:

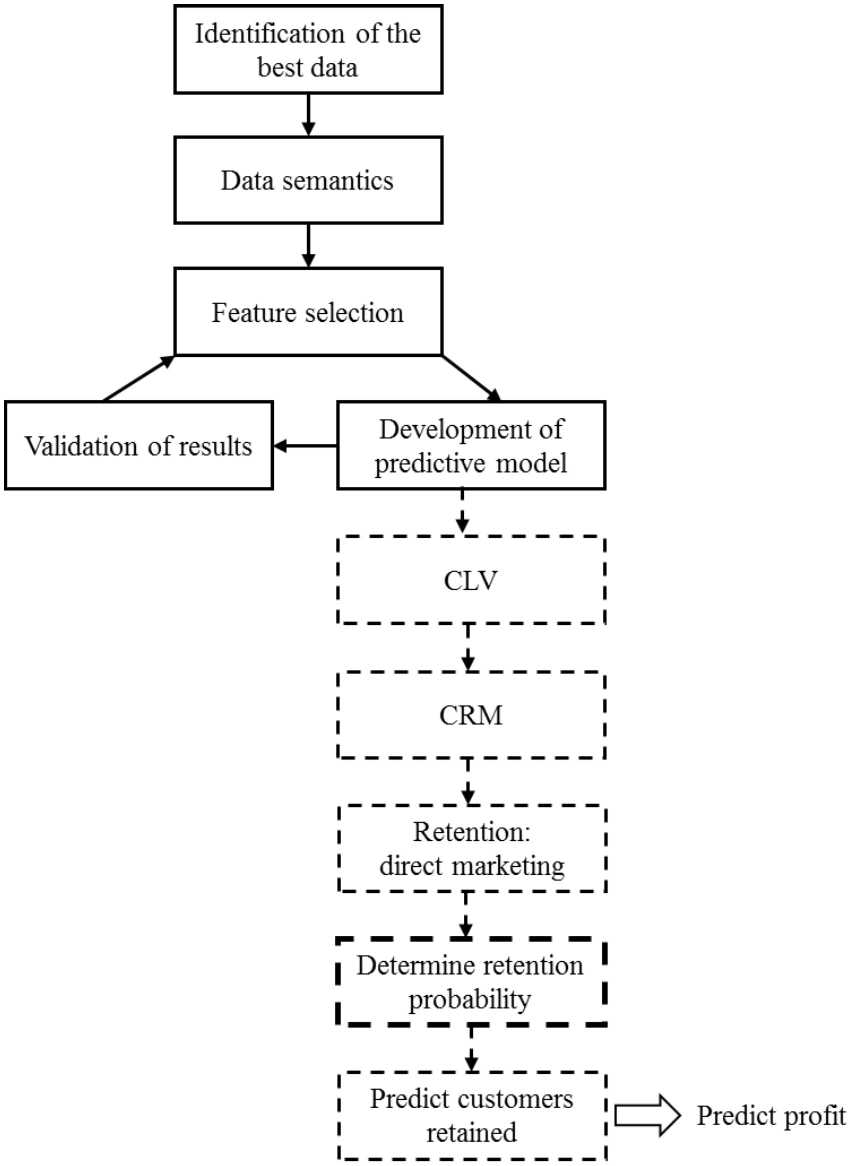

Fig. 8: The Stages of a Predictive 'Return on Data Mining'-Framework
Source: Own development, based on Hadden et al. 2005: 2904.

As can be seen, the single stages mainly remain the same, except for one step that is inserted: to determine retention probability. Practically, this has to follow the step of retention measures, because only after they have been carried out and tested, their success rates and thus the probability of future acceptance can be estimated.

With this retention probability, it would then be possible to estimate, which of the churners is going to be retained. With the use of the CLV-model, their value was determined, hence by combining those two pieces of information, the profit secured can be estimated in advance.

This is an approach which has not been discussed by literature so far.

The idea to combine churn prediction with profitability, however, has generally started to rise, although only very recently, such as in 2012 or 2013. As Kim and Moon write, the "traditional retention management program is strongly tied to the accurate identification of churn candidates and many algorithms have been developed for *that* [emphasis added] purpose" (Kim & Moon 2012: 11719) – but not for the purpose of defining retention rates, or measuring the success of data mining.

For reasons of completeness and inspiration, however, some of the efforts to combine churn management with high customer value are going to be presented briefly and discussed in the following paragraphs.

In February 2013, a work was published that combined churn prediction with the selection of high-value customers (cf. Abbasimehr, Setak & Soroor 2013). They also follow the idea that it is indeed painful to lose customers – but that is more painful losing high-value customers than low-value customers. Consequently, they developed a 'two-phase framework' (cf. ibid.: 1279), which first selects all high value customers and secondly all those with high probability to churn. This is the inverse order compared to the framework suggested above, but it should lead to the same individuals to address. However, the authors do not have the intention to make their gains quantifiable in any way, instead, they again compare different methods for their churn forecasting model: "Results of comparison indicate that the neuro-fuzzy techniques perform better than neural network models and they are a good candidate for churn prediction purposes."

(ibid.: 1279) This is of course of no help for practitioners in so far as it does neither justify the expenses made for data mining, nor give at hand any quantifiable gains achieved with the efforts.

In a paper by Kim and Moon, published in October 2012, it is “claimed that optimal retention management models *should* [emphasis added] consider not only churn probability but also retention probability and expected revenues from target customers” (Kim & Moon 2012: 11718), which is similar to what this work suggested as it concerns the combination of churn, retention and CLV. However, as already indicated by the formulations of ‘claim’ and ‘should’, it was a new approach examined – they set up and compared five different models, and, as quoted earlier, they had problems calculating the retention probability: “Therefore, we concluded that the currently available data sets alone do not contain sufficient information to estimate the retention probability” (Kim & Moon 2012: 11721).

This shows that the idea of putting the pieces together and to work across topics is taking form in data mining research, but that it is still at the brink of break-through.

7. Conclusion

The aim of this work was to investigate the chances of data mining. This was to be done by referring to subscription-based businesses and with the means of literature analysis.

Therefore, in order to set the thematic borders, the first step was to define the topic of subscriptions. This included to outline what the term exactly means, to show the relevance of subscriptions for today’s economy and what can be said about subscriptions from the consumer’s side.

Subsequently followed a chapter defining data mining. This was especially important, because data mining has become a very frequently used phrase, which can mean that it is a term hyped, but not well understood. To avoid such a lack of clarity, data mining was explained starting at its origin, the data warehouse. This set the common ground for further progress.

The next section, number 3, was the consequential combination of those two topics. Churn management is *the* data mining application in the context of subscription-based businesses, primarily for mobile telecommunication companies. It explained the development of predictive models in order to find out in advance about those customers intent to cancel their contract. Because churn is a very special term, it was explained in detail, as well as the concept of churners. Another chapter then explained how churn prediction works and what can be used as explaining variables when creating one.

The following section presented the literature review done on the topic and led to two main findings. Firstly, the discipline revolves very much around the different methods for churn prediction, which means that papers usually apply several techniques in order to compare them. This bears the risk of what was described as meta-focus. Secondly, it was discovered that churn research mainly works with data from the telecommunication industry, which was shown in the second literature analysis. This was referred to as thematic monopolisation, along with the recommendation to expand research to different fields of subscription-based businesses.

Those two discoveries were followed by the development of the framework, which was derived from two main considerations. Firstly, there is the circumstance that churn prediction is so much occupied with itself that it seems to loose strategic focus. Secondly, the argument that it is no advantage for an organisation when churners are known in advance but not prevented. This resulted in the development of a descriptive framework, which is based on churn prediction but organisationally as well as methodically continues its capabilities.

This framework bears two advantages. It increases efficiency, because it makes retention efforts focus on the most profitable customers and neglects those that are less worth retaining. Additionally, it ‘gets things down to one number’, which is the profit that was retained by the data mining activities.

Thus the approach suggests a way how to employ the limited resources of CRM with maximum effectiveness, and can in the end quantify what was referred to as ‘return on data mining’: the beneficial effects of applying data analysis. The sum that stands at the end of the process can serve as a new metric to expand data mining activities within organisations.

An overview over the individual parts explained the causal structure of the framework. Predominantly, this was the concept of Customer Lifetime Value. This approach was used as a means to select the most profitable customers amongst those predicted to churn. Basically, it is another data mining application that has not yet found its general use in the area of subscription-based businesses. This was proven with a third literary analysis, which consequently recommended the increased use of CLV-models in the context of subscription research.

The next chapter subsequently highlighted the relevance of Customer Relationship Management, for it is the department that has the task to run retention measures. Those measures were outlined as well, along with the help of direct marketing and, what should not be neglected in the context of subscriptions, the relevance of switching costs.

From the presentation of the descriptive framework then derived the idea of a predictive model for the same purpose, as a suggestion for forthcoming research.

Key to such a predictive model is the determination of a retention probability so that it can be estimated which of the selected profitable churners is likely to extend their contract when offered appropriate incentives.

To round off the concept of combining churn with CLV and with retention probability, it was shown where similar ideas have appeared in research. As could be seen, the topic started to rise about a year ago, however, much remains to be proven and the concepts have been rather scarce. A framework to calculate the ‘chances of data mining’ for subscription-based companies with the help of a single metric that measures return has not yet been proposed. This highlights the up-to-date-ness as well as the relevance of this work.

To sum it up, the following implications have been found and outlined.

Organisational Implications

Organisations that might see an advantage in churn prediction, but were nevertheless hesitant about investing heavily in data mining - it might make the impression to be a department not generating turnover - could be convinced of the beneficial effects of data mining by the framework proposed. Indeed, "the implementation of such [data mining] models [...] is often subject to practical constraints (e.g. budget constraints)" (Prinzie & Van den Poel 2005: 630). However, the metric suggested renders the chances of data mining palpable.

Research Implications

The first research gap was on churn prediction research. The discipline should apply their models to other industries, especially subscription-based e-commerce or online content providers. This is interesting for the reason that purely online based business models have the advantage of a greater choice of data items, such as "data on customer login behavior and transaction data" (Dover & Murthi 2006: 6), for instance, but also detailed usage times, shopping basket or memory list activities, searches for content or products, for example, or system crashes of a new app, to only name a few.

The second research gap was that Customer Lifetime Value-model research has barely been applied to subscription-based businesses. The mobile communication industry has at least been of moderate interest, probably because there is enough data available and because there is enough (economic) interest into the potential profits of the topic, but working with the concept on industries such as those described above might reveal new valuable insights.

The third research gap was the discovery of lacking findings on retention probability because of this work's suggestion to research on a predictive model developed on basis of the descriptive approach. This model would have to include the probability to accept respective retention measures. Unfortunately, so far no outcomes on acceptance rates are available. A research proposal would therefore be to investigate individual acceptance rates of single marketing retention measures, using this rate to build the predictive framework and to estimate the gains to be expected.

All in all, it can be concluded that this work has successfully portrayed the chances of data mining for subscription-based businesses by analysing relevant literature. Moreover, an independent framework was developed, which makes it possibly to quantify those ‘chances’. In addition to that, two research gaps were identified and a third suggested how to develop the descriptive framework into a predictive model.

At the end a side note: Billy Beane and the Oakland Athletics were the first team in more than hundred years of American-League-Baseball to win twenty games in a row – and that with much less financial means at their disposal to buy talented players than most of their competitors. (cf. Geiselberger & Moorstedt 2013: 9)

Works Cited

Abbasimehr, H., M. Setak & J. Soroor (2013), “A Framework for Identification of High-value Customers by Including Social Network based Variables for Churn Prediction using Neuro-fuzzy Techniques”, in: *International Journal of Production Research*, 51: 4, 1279-1294.

Abrams, D. (April 14, 2013), “Read Petit: A New Serialized Digital Shorts Subscription Service”, Publishing Perspectives, via: http://publishingperspectives.com/2013/04/ read-petit-a-new-serialized-shorts-subscription-service/ (15.11.2013).

Adjei, M., S. Noble & C. Noble (2010), “The Influence of C2C Communications in Online Brand Communities on Customer Purchase Behaviour”, in: *Journal of the Academy of Marketing Science*, 38: 5, 634-653.

Ahn, J.-H., S.-P. Han & Y.-S. Lee (2006), “Customer Churn Analysis: Churn Determinants and Mediation Effects of Partial Defection in the Korean Mobile Telecommunications Service Industry”, in: *Telecommunications Policy*, 30: 10–11, 552–568.

Amazon.com, Inc. (March 28, 2013), *Press Release. Amazon.com to Acquire Goodreads*, via: http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-news Article &ID=1801563&highlight= (16.11.2013).

Benoit, D. & D. Van den Poel (2009), “Benefits of Quantile Regression for the Analysis of Customer Lifetime Value in a Contractual Setting. An Application in Financial Services”, in: *Expert* Systems with Applications, 36: 7, 10475–10484.

Bertoni, S. (September 05, 2013), “Oyster Launches Netflix for Books”, Forbes, via: http://www.forbes.com/sites/stevenbertoni/2013/09/05/oyster-launches-netflix-for-books/ (15.11.2013).

Bhatnagar, A. & S. Ghose (2004), “Segmenting Consumers Based on the Benefits and Risks of Internet Shopping”, in: *Journal of Business Research*, 57, 1352– 1360.

Bhattacherjee, A. (2001), “Understanding Information Systems Continuance: An Expectation-Confirmation Model”, in: *MIS Quarterly*, 25: 3, 351-370.

Blank, G. (2013), “Blurring the Boundaries: New Social Media, New Social Science (NSMNSS)”, in: *International Journal of Market Research*, 55: 3, 461-464.

Bloemer J., T. Brijis, K. Vanhoof & G. Swinnen (2002), “Comparing Complete and Partial Classification for Identifying Customers at Risk”, in: *International Journal of Research in Marketing*, 20, 117–131.

Bolton, R., K. Lemon & P. Verhoef (2004), “The Theoretical Underpinnings of Customer Asset Management. A Framework and Propositions for Future Research”, in: *Journal of the Academy of Marketing Science*, 32: 3, 271-292.

Bose, I. & X. Chen (2009), “Quantitative Models for Direct Marketing. A Review from Systems Perspective”, in: *European Journal of Operational Research*, 195, 1–16.

Braun, M. & D. Schweidel (2010), “Modeling Customer Lifetimes with Multiple Causes of Churn”, via: https://marketing.wharton.upenn.edu/mktg/assets/File/comprisk1.pdf. (12.10.2013).[2]

Carmody, T. (March 29, 2013), “Goodreads is no Instagram: Amazon paid about $150 million”, *The Verge*, http://www.theverge.com/2013/3/29/4161586/goodreads-no-instagram-amazon-paid-150-million (16.11.2013).

Castellanos, M., C. Gupta, S. Wang, U. Dayal, M. Durazo (2012), “A Platform for Situational Awareness in Operational BI”, in: *Decision Support Systems*, 52: 4, 869–883.

Chan, C. (2008), “Intelligent Value-based Customer Segmentation Method for Campaign Management: A Case Study of Automobile Retailer”, in: *Expert Systems with Applications*, 34: 4, 2754–2762.

Chan, S.L., W.H. Ip & V. Cho (2010), “A Model for Predicting Customer Value from Perspectives of Product Attractiveness and Marketing Strategy”, in: *Expert Systems with Applications*, 37: 2, 1207–1215.

Chao, H.-p. (2012), “Competitive Electricity Markets with Consumer Subscription Service in a Smart Grid”, in: *Journal of Regulatory Economics*, 41: 1, 155–180.

Chordas, L. (2001), “Building a better Warehouse”, in: *Best's Review*, 101: 11, 117-121.

Clemons, E. K. & G. Gao (2008), “Consumer Informedness and Diverse Consumer Purchasing Behaviors: Traditional Mass-market, Trading down, and Trading out into the Long Tail”, in: *Electronic Commerce Research and Applications*, 7: 1, 3–17.

Conyette, M. (2011), “Demographics for Segmentation in Online Travel”, in: *International Journal of Trade, Economics and Finance*, 2: 1, 93-98.

[2] Also available in: Braun, Michael & David Schweidel (2011), “Modeling Customer Lifetimes with Multiple Causes of Churn”, in: *Marketing Science*, 30: 5, 881-902.

Coussement, K. & K. W. De Bock (2013), “Customer Churn Prediction in the Online Gambling Industry: The Beneficial Effect of Ensemble Learning”, in: *Journal of Business Research*, 66: 9, 1629–1636.

Coussement, K. & D. Van den Poel (2008), “Churn Prediction in Subscription Services: An Application of Support Vector Machines while Comparing two Parameter-selection Techniques”, in: *Expert Systems with Applications*, 34: 1, 313–327.

Coussement, K. & D. Van den Poel (2009), “Improving Customer Attrition Prediction by Integrating Emotions from Client/Company Interaction Emails and Evaluating Multiple Classifiers”, in: *Expert Systems with Applications*, 36: 3, Part 2, 6127–6134.

Datta, P., B. Masand, D. Mani & B. Li (2001), “Automated Cellular Modeling and Prediction on a Large Scale”, in: *Artificial Intelligence Review*, 14: 6, 485–502.

dbw (November 23, 2012), “booxl, Another Spotify for Ebooks Slated to Launch in Beta this February”, Digital Book World, via: http://www.digitalbookworld.com/ 2012/booxl-another-spotify-for-ebooks-slated-to-launch-in-beta-this-february/ (15.11.2013).

De Bock, K. & D. Van den Poel (2012), “Reconciling Performance and Interpretability in Customer Churn Prediction using Ensemble Learning based on Generalized Additive Models”, in: *Expert Systems with Applications*, 39: 8, 6816–6826.

Dover, H. & B. Murthi (2006), “Asymmetric Effects of Dynamic Usage Behavior on Duration in Subscription-based Online Services”, in: *Journal of Interactive Marketing*, 20: 3-4, 5-15.

Fruchter, G. & S. Sigué (2013), “Dynamic Pricing for Subscription Services”, in: *Journal of Economic Dynamics and Control*, 37: 11, 2180–2194.

Geiselberger, H. & Moorstedt, T. (2013), *Big Data. Das neue Versprechen der Allwissenheit*, Berlin: Suhrkamp Verlag.

Glady, N., B. Baesens & C. Croux (2009), “Modeling Shurn Using Customer Lifetime Value”, in: *European Journal of Operational Research*, 197: 1, 402–411.

Glady, N., B. Baesens & C. Croux (2009a), “A Modified Pareto/NBD Approach for Predicting Customer Lifetime Value”, in: *Expert Systems with Applications*, 36: 2, Part 1, 2062–2071.

Goodreads (2013), *“Once upon a time …” is much more exciting than “The End”*, via:

https://www.goodreads.com/jobs?anchor_added=true&jvi=oI84XfwB%2CJob#openpositions (16.11.2013).

Gupta, S. & D. R. Lehmann (2003), “Customers as Assets”, in: *Journal of Interactive Marketing*, 17: 1, 9–24.

Hadden, J., A. Tiwari, R. Roy & D. Ruta (2005), “Computer Assisted Customer Churn Management: State-of-the-art and Future Trends”, in: *Computers & Operations Research*, 34: 10 (2007; published online 2005), 2902–2917, via: http://dx.doi.org/10.1016/j.cor.2005.11.007 (10.11.2013).

Haenlein, M. (2013), “Social Interactions in Customer Churn Decisions: The impact of Relationship Directionality”, in: *International Journal of Research in Marketing*, 30: 3, 236–248.

Haenlein, M., A. Kaplan & A. Beeser (2007), “A Model to Determine Customer Lifetime Value in a Retail Banking Context”, in: *European Management Journal*, 25: 3, 221–234.

Han, S. H., S. X. Lu & S. Leung (2012), “Segmentation of Telecom Customers based on Customer Value by Decision Tree Model”, in: *Expert Systems with Applications*, 39: 4, 3964–3973.

Hashmi, N., N. Butt & M. Iqbal (2013), “Customer Churn Prediction in Telecommunication. A Decade Review and Classification”, forthcoming in: *International Journal of Computer Science Issues*, 10: 5, available online via: http://www.researchgate.net/publication/257920014_Customer_Churn_Prediction_in_Telecommunication_A_Decade_Review_and_Classification/file/e0b495261475ba6767.pdf (26.11.2013).

Hiziroglu, A. & S. Sengul (2012), “Investigating Two Customer Lifetime Value Models from Segmentation Perspective”, in: Procedia - Social and Behavioral Sciences, 62, 24 October 2012, 766–774.

Huang, B., M. Kechadi & B. Buckley (2012), “Customer Churn Prediction in Telecommunications”, in: *Expert Systems with Applications*, 39: 1, 1414–1425.

Huang, B., T.-M. Kechadi, B. Buckley, G. Kiernan, E. Keogh & T. Rashid, (2010), “A new Feature Set with new Window Techniques for Customer Churn Prediction in Land-line Telecommunications”, in: *Expert Systems with Applications*, 37: 5, 3657–3665.

Huang, P., N. Lurie & S. Mitra (2009), “Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods”, in: *Journal of Marketing*, 73, 55–69.

Huffier (2013), via: http://www.huffier.com/ (20.11.2013).

Hughes, J. (2010), “Supplying Web 2.0: An Empirical Investigation of the Drivers of Consumer Transmutation of Culture-oriented Digital Information Goods”, in: *Electronic Commerce Research and Applications*, 9, 418–434.

Hung, S.-Y., D. Yen & H.-Y. Wang (2006), “Applying Data Mining to Telecom Churn Management”, in: *Expert Systems with Applications*, 31: 3, 515–524.

Ho, H.-Y., L.-W. Wang & H.-J. Cheng (2011), “Authors, Publishers, and Readers in Publishing Supply Chain: The Contingency Model of Digital Contents Production, Distribution, and Consumption”, *Systems Engineering Procedia*, 2, 398 – 405.

Hwang, H., T. Jung & E. Suh (2004), “An LTV Model and Customer Segmentation based on Customer Value. A case Study on the Wireless Telecommunication Industry”, in: *Expert Systems with Applications*, 26: 2, 181–188.

Jahanzeb, S. & S. Jabeen (2007), “Churn Management in the Telecom Industry of Pakistan: A Comparative Study of Ufone and Telenor”, in: *Database Marketing & Customer Strategy Management*, 14: 2, 120–129.

Karahoca, A. & D. Karahoca (2011), “GSM Churn Management by using fuzzy c-means Clustering and Adaptive Neuro fuzzy Inference System”, in: *Expert Systems with Applications*, 38: 3, 1814–1822.

Keaveney, S. & M. Parthasarathy (2001), “Customer Switching Behavior in Online Services: An Exploratory Study of the Role of Selected Attitudinal, Behavioral, and Demographic factors”, in: *Journal of the Academy of Marketing Science*, 29: 4, 374-390.

Khajvand, M., K. Zolfaghar, S. Ashoori & S. Alizadeh (2011), “Estimating Customer Lifetime Value based on RFM Analysis of Customer Purchase Behavior: Case Study”, in: *Procedia Computer Science*, 3, 57–63.

Kim, S.-Y., T-S. Jung, E.-H. Suh & H.-S. Hwang (2006), “Customer Segmentation and Strategy Development based on Customer Life Time Value: A Case Study”, in: *Expert Systems with Applications*, 31: 1, 101–107.

Kim, Y. S., H. Lee & J. D. Johnson (2013), “Churn Management Optimization with Controllable Marketing Variables and Associated Management Costs”, in: *Expert Systems with Applications*, 40: 6, 2198–2207.

Kim, Y. S. & S. Moon (2012), “Measuring the Success of Retention Management Models built on Churn Probability, Retention Probability, and Expected Yearly Revenues”, in: *Expert Systems with Applications*, 39: 14, 11718-11727.

Kim, M.-K., M.-C. Park & D.-H. Jeong (2004), “The Effects of Customer Satisfaction and Switching Barrier on Customer Loyalty in Korean Mobile Telecommunications Services”, in: *Telecommunications Policy*, 28: 2, 145-159.

Lam, S., V. Shankar, M. K. Erramilli & B. Murthy (2004), “Customer Value, Satisfaction, Loyalty, and Switching Costs: An Illustration From a Business-to-Business Service Context”, in: *Journal of the Academy of Marketing Science*, 32: 3, 293-311.

Lemmens, A. & C. Croux (2006), “Bagging and Boosting Classification Trees to Predict Churn”, in: *Journal of Marketing Research*, 43: 2, 276-286.

Liao, S.-H., P.-H. Chu & P.-Y. Hsiao (2012), “Data Mining Techniques and Applications. A Decade Review from 2000 to 2011”, in: *Expert Systems with Applications*, 39: 12, 11303–11311.

Liu, M., X.-q. Qiao & W.-l. Xu (2011), "Three Categories Customer Churn Prediction Based on the Adjusted Real Adaboost", in: *Communications In Statistics: Simulation & Computation*, 40: 10, 1548-1562.

Min, D. & L. Wan (2009), “Switching Factors of Mobile Customers in Korea”, in: *Journal of Service Science*, 1: 1, 105-120.

Neslin, S., S. Gupta, W. Kamakura, J. Lu & C. Mason (2006), “Defection Detection: Measuring and Understanding the Predictive Accuracy of Customer Churn Models”, in: *Journal of Marketing Research*, 43: 2, 204-211.

Ngai, E., L. Xiu, & D. Chau (2009), “Application of Data Mining Techniques in Customer Relationship Management: A Literature Review and Classification”, in: *Expert Systems with Applications*, 36: 2, Part 2, 2592–2602.

Olanoff, D. (March 28, 2013), “Amazon Acquires Social Reading Site Goodreads, Which Gives The Company A Social Advantage Over Apple”, TechChrunch.com, http://techcrunch.com/2013/03/28/amazon-acquires-social-reading-site-goodreads/ (16.11.2013).

Oysterbooks.com, via: www.oysterbooks.com (20.11.2013).

Phadke, C., H. Uzunalioglu, V. Mendiratta, D. Kushnir, & D. Doran (2013), “Prediction of Subscriber Churn Using Social Network Analysis”, in: *Bell Labs Technical Journal*, 17: 4, 63-75.

Prinzie, A. & D. Van den Poel (2005), “Constrained Optimization of Data-mining Problems to Improve Model Performance: A Direct-marketing Application”, in: *Expert Systems with Applications*, 29: 3, 630–640.

Rainardi, V. (2011), *Building a Data Warehouse: Examples in SQL Server*, Berkeley: Apress.

Samimi, A. & A. Aghaie (2011), “Using Logistic Regression Formulation to Monitor Heterogeneous Usage Rate for Subscription-based Services”, in: *Computers & Industrial Engineering*, 60, 89-98.

Shy, O. (2008), “Measuring the Cost of Making Payment Decisions”, in: *The Journal of Socio-Economics*, 37, 2411–2416.

Stewart, D., M. Hess & H. Nelder (2011), “How Relevancy, Use, and Impact Can Inform Decision Making. The Uses of Quantitative Research”, in: *Supplement Journal of Advertising Research*, 51, 195-206.

Sun Z., G. Bebis & R. Miller (2004) “Object Detection Using Feature Subset Selection”, in: *Pattern Recognition*, 37, 2165–2176.

Thohira, M., M. Chambers & N. Sprague (2010), “Full-Text Databases: A Case Study Revisited a Decade Later”, in: *Serials Review*, 36: 3, 152‐160.

Thomas, Owen (March 29, 2013), “People are Howling about the $1 Billion Price BusinessWeek made up For Amazon's Goodreads Acquisition”, *Business Insider*, via: http://www.businessinsider.com/amazon-goodreads-purchase-price-2013-3 (16.11.2013).

Verbeke, W., K. Dejaeger, D. Martens, J. Hur & B. Baesens (2012), “New Insights into Churn Prediction in the Telecommunication Sector: A Profit Driven Data Mining Approach”, in: *European Journal of Operational Research*, 218: 1, 211–229.

Verbeke, W., D. Martens, C. Mues & B. Baesens (2011), “Building Comprehensible Customer Churn Prediction Models with Advanced Rule Induction Techniques”, in: *Expert Systems with Applications*, 38: 3, 2354–2364.

Wasserman, Steve (2012), “The Amazon Effect”, in: *The Nation*, June 18 (2012), 13-22.

Wong, K. K.-K. (2011), “Using Cox Regression to Model Customer Time to Churn in the Wireless Telecommunications Industry”, in: *Journal of Targeting, Measurement and Analysis for Marketing*, 19: 1, 37-43.

Declaration of Authorship

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Mastersrarbeit mit dem Thema: “The Chances of Data Mining in Subscription-based Businesses: A Literature-based Analysis” selbstständig und ohne fremde Hilfe angefertigt habe.
Die Übernahme wörtlicher Zitate sowie die Verwendung der Gedanken anderer Autoren habe ich an den entsprechenden Stellen der Arbeit kenntlich gemacht.
Ich bin mir bewusst, dass eine falsche Erklärung rechtliche Folgen haben wird.

Friedrichshafen, 05.12.2013	Kerstin Weinheimer
Ort, Datum	**Unterschrift**