# Applying client churn prediction modeling on home-based care services industry

Raul Manongdo
Advanced Analytics Institute
University of Technology
Sydney, Australia
Email: raul.manongdo@student.uts.edu.au

Guandong Xu
School of Software and
Advanced Analytics Institute
University of Technology
Sydney, Australia
Email: guandong.xu@uts.edu.au

***Abstract*—Client churn prediction model is widely acknowledged as an effective way of realizing customer life-time value especially in service-oriented industries and under a competitive business environment. Churn model allows targeting of clients for retention campaigns and is a critical component of customer relationship management(CRM) and business intelligence systems.**

**There are numerous statistical models and techniques applied successfully on data mining projects for various industries. While there is literature for prediction modeling on hospital health care services, non-exist for home-based care services. In this study, logistic regression, random forest and C5.0 decision tree were the models used in building a binary client churn classifier for a home-based care services company based in Australia.**

**All models yielded prediction accuracies over 90% with tree based classifiers marginally higher and C5.0 model found to be suitable for use in this industry. This study also showed that existing client satisfaction measures currently in use by the company does not adequately contribute to churn analysis.**

## I. Introduction

Client churn prediction is one of the earlier application of data mining widely used in large companies with acknowledged benefits. In telecommunication industry, it is claimed that it costs 3 times more to acquire a new customer than to retain existing ones. Churn model allows targeting of clients for retention campaigns and hence, investment on retention models and campaigns yields better returns [3].

Many kinds of literature exist about successful client churn prediction projects. While there are for hospital-based and health care services, the author was unable to find application for home-based care services. The objective of the study therefore is to develop an initial client churn prediction model in the context of an actual home-based care services company, present the challenges encountered and propose steps to further develop and introduce client churn and retention models.

Home-based care services are in many respects different from acute care, office-based health or mental health care [13].

- The services are performed at the consumers daily living situation and is frequently for a long and extended time period. Clients are mostly frail and disabled wherein service delivery expectations, preferences and issues may not be clearly and objectively communicated.
- In general, services are varied and low-tech, often provided by personnel with limited training and without professionally derived standards of practice.
- Clients have strong opinions and preferences on services received having performed the same when younger and able-bodied.
- Industry is labor intensive with high client-facing to non-facing work hours by staffs.
- In general, service providers are small to medium sized companies with lesser years in operation and lower level of IT maturity and capability as compared to large and established companies.

Many statistical models had been successfully used as binary churn classifiers. This includes logistic regression, decision trees, boosting algorithms (e.g. variants of AdaBoost), boosted trees (gradient boosted decision trees), random forest, neural networks, evolutionary computation (e.g. genetic algorithm and ant colony optimization), an ensemble of support vector machines and an ensemble of hybrid methods [3].

In this study, we used three statistical models; logistic regression (GLM), random forest (RF) and C5.0 decision tree. Logistic regression was chosen as base model as it is easy to interpret, robust and widely used in the industry. Random forest is an ensemble model claimed to have better prediction accuracies because of combining outputs from weak classifiers and random selection of observations and features [7]. C5.0 also combines outputs from weak classifiers through boosting of misclassified predictions for each successive tree iteration.

The plan of this article is to define client churn, metrics and monitoring as formulated for this case company. We then describe the data preparation stage and present client churn and associated client satisfaction charts. We proceed with methodology and how it was performed. We then discussed similar client churn experiments and projects from literature surveyed. The conclusion then summarizes the study and possible steps moving forward.

## II. Problem Statement

The goal is to develop an initial client churn prediction model for a home-based care services company in particular and to the industry in general. In the process, challenges encountered will be presented in this paper. The prediction model that exhibits the highest prediction accuracy that can readily be adapted for use in retention campaigns will be selected. An associated client satisfaction monitoring system as it relates to client churns will also be reviewed. The aim of this paper is not to justify in detail a selection of a particular model but to develop an initial working model that can readily be adopted and for subsequent enhancement.

### A. Churn Definition

A churned client is an existing customer who discontinued enrollment to all of a companys services due to an unsatisfactory experience or unfavorable perception. A *Discharge Reason* specifies this customer-supplied reason and not all reasons are churn related. For example, *Dissatisfied with company* is considered while *Change in living arrangement* is not.

Some programs/services are not monitored for churns. Services may have fix expiration terms or are one-off incident in nature. For example, the government assisted transitional care from a hospital is fixed for 12 weeks and will terminate regardless. Ambulance service is on-call basis and ends after the service is performed. To differentiate, services are tagged as **Core** if monitored and **Non-core** if not. A client can enroll to multiple services at the same time which can be a combination of core and non-core services. Multiple services are often bundled into a single program and in this paper, the terms services and programs are used interchangeably.

$$Churn\,Flag \Leftarrow function$$
$$(Client\,Program\,Enrollments, Discharge\,Reason)$$
$$\Leftarrow \begin{cases} 1\ (\textbf{Churn})\,Discharge\,Reason \in R\ and \\ \quad Client\,Program\,Enrollment\ type \in P\ and \\ \quad Client\,Program\,Enrollment\ count = 0 \\ 0\ (\textbf{non Churn})\,otherwise. \end{cases} \tag{1}$$

where $R$ is a set of Discharge Reasons $r_1, r_2..., r_n$ and $P$ is a set of 'Core' Client Programs $p_1, p_2..., p_n$ and count is the remaining enrolled program(s) after discharge.

The count of client churns is monitored over a period of time. On a monthly basis, the formula is

$$\begin{aligned} Active\ Clients_{(at\ \ end\,month)} = & \\ & \sum Active\ Clients_{(at\ \ begin\,month)} \\ & + \sum new\ Clients_{(within\,month)} \\ & - \sum Churned\ Clients_{(within\,month)} \\ & - \sum Non\,churn\ Client\ Exits_{(within\,month)} \end{aligned} \tag{2}$$

A Churn rate is a percentage of churned clients over active clients covering a geographical state within a given time period. Figure 2 expressed the churn rate annually and Figure 3 monthly.

$$Churn\,Rate\% = \sum_{i=1}^{k} \frac{Churned\ Clients^{k}_{(within\ \ month)}}{Active\ Clients^{k}_{(at\ \ end\,month)}} \tag{3}$$

where K is set of States $s_1, s_2.., s_k$

### B. Methodology

In this supervised data mining project, data was sourced from business support (i.e. client satisfaction surveys) and operational support systems (i.e. data mart). We selected attributes pertaining to client demographics, client-expressed home care needs and service preferences. From operational customer data, we obtained a history of program enrollments and discharges, service delivery, complaints, interactions and satisfaction surveys.

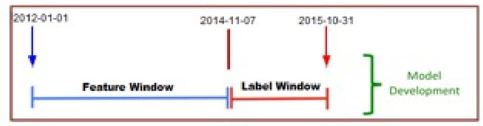



Fig. 1. Model development windows

Prior to the start of the feature window i.e. (1-Jan-2012), inactive clients were filtered out of the dataset. Within the feature window, a clients historical data were used as attributes for feature selection. Attributes were crafted and derived from base variables such as a *count of client issues*, *highest complaint escalation level reached*. For the label window (i.e. after 07- Nov-2015), clients were labeled as churned or non-churned as defined in the previous section and recorded at the time the churn occurred. The intention is to capture sufficient historical and empirical data for 2 years prior to observing client churns the following year.

For feature selection, step-wise GLM and correlation analysis were used with a few important variables as identified from RF added. For binary prediction modeling, logistic regression, random forest and C5.0 decision tree were attempted.

To create a training set, stratified sampling of observations was used. 80% of the population was set aside for training and 20% for testing. Stratified sampling guaranteed both sets have an equal percentage of churns and non-churn clients for training and testing. 5-fold cross-validation(CV) and ROC curves will be used for further model comparison . We chose 5 as this is the minimum number of folds that ensured every observation will be included in the CV test sets given that 20% of the dataset is for testing.

In terms of tools used, R Language and WEKA (Waikato Environment for Knowledge Analysis) were used for prediction modeling and SQL Server, Excel and R for data collection and analysis.

## III. Case Study: Churn Prediction in home-based care services company

Prerequisite tasks were needed to be defined and established prior to any data analysis.

- Define and formulate client churns. The definition is as described in the previous section.
- De-duplication of client IDs. It was found that many client person was identified by multiple client IDs due to the inconsistent recording of client names and unrelated multiple client program enrollments. To establish

uniqueness, a clients date of birth and full name were used with client naming conventions standardized.

In so doing, the client accounts base was reduced to 18%.

### A. Dataset collection and preparation

There were significant missing values and outliers and the following data cleansing steps were undertaken:

- Variables with more than 50% missing values were dropped from consideration
- Mean values imputed for missing numeric values
- Median values were set for numeric outliers (i.e. 3 times the interquartile range over and below the 1st and 3rd quartile value)
- Median values or sensible defaults were set for missing and outlying categorical values

The cleansing exercise reduced the dimension of the dataset to 12,526 instances by 17 features.

There were many restrictions that hindered data collection and analysis. Critical attributes such as *Client Status* (i.e. *Active* or *Discharged*) or alternate sources (e.g. *Discharge Date*) were not consistently populated and hence unreliable. As an alternative, we defined *Active Client* as a customer having a recorded timesheet in the future. Important attributes such as *Age* can not be derived from *Date of Birth* was missing in more than half of the dataset. Similarly, *Gender* had 38% missing values. Other data like *Discharge Reasons* were only captured just a year earlier hence restricting the available time windows for development as shown in Figure 1.

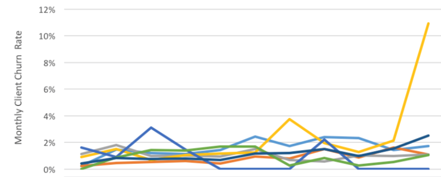
### B. Descriptive Analysis

Below are graphs showing churn ratios and trends in client churns and customer satisfaction. It is intended to show the utility of monitoring and the relationship of churns to client satisfaction as currently practiced. The graphs are composed at the state level as the customer relationship management job function is owned by state managers.


| | State 1 | State 2 | State 3 | State 4 | State 5 | State 6 | National |
|---|---|---|---|---|---|---|---|
| Active Clients as of Oct '15 | 408 | 4,813 | 908 | 1,116 | 125 | 379 | 7,749 |
| Churned Clients between Dec-'14 to Oct'15 | 76 | 427 | 87 | 287 | 12 | 36 | 925 |
| Annual Churn Rate | 18.6% | 8.9% | 9.6% | 25.7% | 9.6% | 9.5% | 11.9% |

Fig. 2. Client churn rate

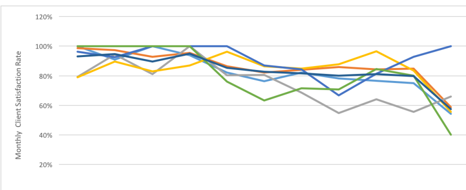
At the national level, Fig. 2 shows that 11.9% of active clients churned between Dec 2014 to Oct 2015. In churn prediction, the positive class (i.e. churned clients) usually represents a small percentage of the population [2]. Notice that distribution of clients across states are abnormal with *State 2* having around 60% of the population reaffirming the state where the company started and became established.




| | Dec-14 | Jan-15 | Feb-15 | Mar-15 | Apr-15 | May-15 | Jun-15 | Jul-15 | Aug-15 | Sep-15 | Oct-15 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| State 1 | 0.2% | 1.5% | 1.2% | 1.1% | 1.4% | 2.4% | 1.7% | 2.4% | 2.3% | 1.4% | 1.7% |
| State 2 | 0.2% | 0.5% | 0.5% | 0.6% | 0.4% | 0.9% | 0.8% | 1.5% | 0.9% | 1.6% | 1.1% |
| State 3 | 1.1% | 1.8% | 1.0% | 1.0% | 0.9% | 1.5% | 0.7% | 0.6% | 1.0% | 1.0% | 1.1% |
| State 4 | 0.9% | 1.5% | 0.7% | 1.0% | 1.2% | 1.3% | 3.7% | 2.0% | 1.3% | 2.1% | 10.9% |
| State 5 | 1.6% | 0.9% | 3.1% | 1.4% | 0.0% | 0.0% | 0.0% | 2.2% | 0.0% | 0.0% | 0.0% |
| State 6 | 0.0% | 0.9% | 1.4% | 1.4% | 1.7% | 1.7% | 0.3% | 0.8% | 0.3% | 0.5% | 1.1% |
| National | 0.4% | 0.8% | 0.7% | 0.8% | 0.7% | 1.2% | 1.2% | 1.5% | 1.0% | 1.5% | 2.5% |

Fig. 3. Percentage Trend of Churned Clients by State

In Figure 3, *State 4* experienced unusually high churn rates for the months of Oct and Nov 2015. On further investigation, the company revealed that during these months, there was no assigned state manager and a replacement took some time. Further drill down analysis, not shown in this paper, identified the discharge reasons most common (i.e. *Quality of Care Received*, *Continuity of Access to Workers*).As for enrolled programs/services, churns occurred mostly in programs and services under the industry competitive *Private/Commercial* programs and the government assisted *Home support packages*.




| | Dec-14 | Jan-15 | Feb-15 | Mar-15 | Apr-15 | May-15 | Jun-15 | Jul-15 | Aug-15 | Sep-15 | Oct-15 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| State 1 | 100.0% | 90.9% | 100.0% | 93.8% | 82.1% | 76.3% | 82.2% | 78.2% | 76.7% | 75.0% | 54.2% |
| State 2 | 98.7% | 97.4% | 92.7% | 95.4% | 86.5% | 82.2% | 83.9% | 85.8% | 84.1% | 84.7% | 58.8% |
| State 3 | 79.2% | 94.4% | 81.0% | 100.0% | 80.4% | 80.5% | 68.6% | 54.7% | 64.0% | 55.6% | 65.8% |
| State 4 | 78.9% | 89.5% | 82.9% | 87.0% | 96.2% | 86.4% | 84.9% | 87.6% | 96.4% | 83.3% | 55.6% |
| State 5 | 96.3% | 92.6% | 100.0% | 100.0% | 100.0% | 87.0% | 84.4% | 66.7% | 81.3% | 92.9% | 100.0% |
| State 6 | 100.0% | 100.0% | 100.0% | 100.0% | 76.0% | 63.2% | 71.4% | 70.6% | 84.6% | 80.0% | 40.0% |
| National | 93.1% | 94.7% | 89.6% | 94.8% | 85.4% | 82.7% | 81.6% | 80.0% | 81.1% | 79.8% | 57.4% |

Fig. 4. Percentage Trend of Satisfied Clients by State

In composing the graph in Figure 4, satisfaction with a service is measured by a favorable response from the client to the survey question How likely are you to recommend the company to your family and friends?. It is important to note that the survey sample represents the average only 0.5% of the population. In addition, some states recorded a 100% satisfaction rating which cast doubt on the diligence in capturing data and quality of responses obtained. The monthly client satisfaction survey is composed of 10 questions. None were significantly related to client churn and was not included in selected features.

Churns are inversely related to client satisfaction. Nationally, the average monthly churn to satisfaction ratio is 0.1014; that is, a unit percentage change in churn ratio is reflected by .1014 change in client satisfaction with a standard deviation is 7.69. This excludes the month of October which is an exception.

TABLE I
MONTHLY CHANGES TO CHURN AND SATISFACTION PERCENTAGES

| | Churn % Δ | Satisfaction % Δ | Churn / Satisfaction |
|---|---|---|---|
| Jan-15 | -0.4 | 1.6 | -0.25 |
| Feb-15 | 0.1 | -5.1 | -0.019607843 |
| Mar-15 | -0.1 | 5.2 | -0.019230769 |
| Apr-15 | 0.1 | -9.4 | -0.010638298 |
| May-15 | -0.5 | -2.7 | 0.185185185 |
| Jun-15 | 0 | -1.1 | 0.0 |
| Jul-15 | -0.3 | -1.6 | 0.1875 |
| Aug-15 | 0.5 | 1.1 | 0.454545455 |
| Sep-15 | -0.5 | -1.3 | 0.384615385 |
| Oct-15 | -1 | -22.4 | 0.044642857 |

A rate change in one direction should ideally be associated with an opposite change in the other rate. (i.e.negative Churn/Satisfaction value). There are several months that this is not the case. In addition, the standard deviation of 7.69 is large. Therefore, the client satisfaction measure does not adequately reflect changes in churn rates and at best, may contribute to client retention instead.

In [13], it is claimed that a single item global rating measure as used here is not adequate in capturing the complexity of services in health, mental health or long-term care and an alternative survey method that measured satisfaction over more service delivery categories were proposed.

### C. Feature Selection

Candidate variables for feature selection were primarily obtained from step-wise regression and correlation. A few variables were added as identified by RF. Logistic regression identifies significant variable by the p-value of less than .10. RF ranks variable importance by information gain and Gini index. For C5.0, attributes ranked by is shown below.

1) Client type
2) Average core service hours duration
3) Most used service billing grade
4) Most used home care workers pay category grade
5) Highest complaint escalated level by client
6) Count of client initiated service cancellations
7) Count of issues raised by client
8) Maximum core service duration
9) Minimum core hours duration
10) Client home state
11) Enrolled core programs/services
12) Count of enrolled programs/services

Each feature can be classified based on time period. With the exception of *Client type* and *Client home state*, remaining features were all time-based, aggregated over the feature window.

In direct marketing literature, marketing models had been developed using RFM schemes for more than 30 years [11]. In RFM, customers prior purchases are grouped based on three categories; *Recency*, *Frequency* and *Monetary Value* and are used as indicators for future actions. The RFM model was extended for use in attrition modeling (eRFM) to include socio-demographic client information [13]. Comparing eRFM to our selected features, the 2 non-time based features (i.e (*Client type*,*Client home state*) fall under demographics. Recency (e.g. *elapsed time since the last service*) and frequency (e.g. *average service duration*) also exist. Monetary attributes however are missing and are actually not included in the dataset. In [9], it was however found that in RFM, socio-demographic variables do not play important roles in explaining churn. In addition, it also claimed that important churn predictors belong to the category of variables describing the subscription and client interactions (i.e. *complaints issues*, *service cancellations*).

### D. Model Building

The resulting regression equation from the GLM model shown below had class membership set at 50%. The ROC curve shown in Figure 6 confirmed that this threshold is appropriate and can be used for all candidate prediction models.

$$
\begin{aligned}
GLM\ Client\,Churn\,Probability = -1.842 \\
-0.8587 * (HOME_STR_state =' State5') \\
+0.4012 * (Sex =' Female') \\
+0.4279 * (Sex =' Male') \\
-0.8938 * (ClientType =' Com') \\
-1.289 * (ClientType =' Dis') \\
+0.9334 * (ClientType =' Soc') \\
+2.703 * (ClientType =' You') \\
-0.2915 * (CoreProgramsNums) \\
-0.2012 * (MaxCoreProgramHours) \\
+0.1302 * (MinCoreProgramHours) \\
-0.0074 * (AverageCoreProgramHours) \\
+0.2758 * (MostUsedBillingGrade =' BGrade2') \\
+1.526 * (MostUsedBillingGrade =' BGrade5') \\
+1.076 * (MostUsedBillingGrade =' BGrade6') \\
+2.119 * (complainttier =' Tier1') \\
+1.755 * (complainttier =' Tier2') \\
+2.774 * (complainttier =' Tier3') \\
+0.359 * (Issues_Raised) \\
-0.1433 * (Client_Initiated_Cancellations)
\end{aligned}
$$

For Random Forest, we followed the suggestion to use a large number of trees (i.e. ntree= 1,000) and to set the fix number of randomly selected features in constructing tree nodes (i.e. mtry =3) to equal the square root of the total number of variables [14]. The suggested value of mtry is the same as our generated value using RF package method.

RF does not create a single decision tree from the forest of trees it generates Due to the random sampling of features(i.e. mtry), each tree can only be considered a domain expert of its own sampled data and no single tree can be used as a general classifier for the whole dataset. Unlike RF, C5.0 can generate a single merged tree. The model uses a boosting technique that creates several iterations of the model with varying misclassification weights and the model parameter used was trials = 10.

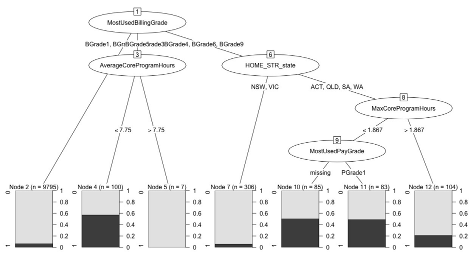



Fig. 5. C5.0 Model Decision Tree

### E. Model Results and Evaluation

The criteria used for comparing prediction models are area under the Receiver Operating Characteristic (ROC) curve and F-score. The area under ROC curve measures the performance of a binary classifier as its discrimination threshold is varied independent of classification cost or class distribution context [10]. F-score is measure of prediction accuracy and can be interpreted as a weighted average of the precision and recall. In model evaluation for this study, we used a variant, $F_2$ which weighs recall twice more than precision; that is by placing more emphasis on false negatives [16].

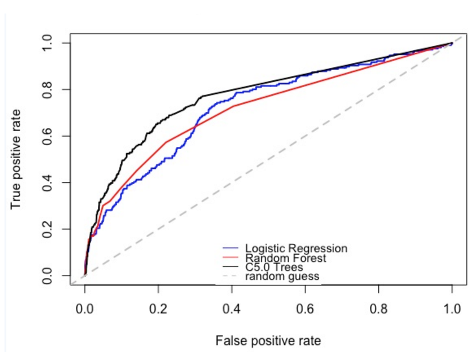



Fig. 6. ROC Curve of Prediction Models

TABLE II
COMPARISON OF MODEL PERFORMANCE

| | $F_2$ Score | Area under ROC | 5-Fold CV Accuracy |
|---|---|---|---|
| Logistic Regression | 1.4197 | 72.89% | 92.099% |
| Random Forest | 1.3939 | 69.82% | 92.168% |
| C5.0 Decision Tree | 1.6864 | 76.03% | 92.138% |

The prediction accuracies marginally differ with logistic regression understandably close as common features were selected using GLM. C5.0 model yielded the highest $F_2$ score, the highest area under ROC curve and marginally lower 5-fold CV prediction accuracy than RF. C5.0 model was also chosen because of ease of understanding decision trees and rule sets due to the familiar if-then-else structure of tree based classifiers. Operationally, the case company utilizes a business rule engine for manpower resource planning and scheduling. The rules set from C5.0 model can, therefore, be readily adapted to the language syntax of the business rule engine without incurring significant cost and effort.

## IV. RELATED WORK

In hospital-based health care services, a binary prediction model of patient readmission was described in [1]. Comparing variables in our dataset, a few domestic services (i.e.*needing help in bathing*) and demographic data (i.e. *Age, Sex, Care Giver*) were found to be common. Their study also included attributes that can be considered in churn prediction; *availability of an informal caregiver, hospital admission rate, not working status, high social deprivation, patient or family education, religious participation, doctor visits and emergency hospital admission within a period, self-rated generated health, hospital stay duration*.

Because clients in home-based care are mostly frail and aged, emotionality indicators can perhaps also be considered. In [7], expressed client emotions in eMails were include as churn predictors. A primary finding of the research is a significant relationship between positively expressed emotions with clients churning.

From the perspective of client satisfaction, studies showed that home-based care satisfaction was not related to *gender* or *ethnicity*, negatively associated with *physical disability* and positively related with *age*. Several attributes on service recovery activities were covered pertaining to remedial actions arising from a customer complaint about service failures [13]. In our dataset, no attribute on service recovery was included.

In [4], a major determinant for client retention in the telecommunication services is the *duration of the client relationship* and *usage of telecommunications services*. Client retention is in general directly related to client satisfaction and a retention model is actually a separate but closely associated with churn models. Retention model considers other factors such as retention promotions and the likelihood of acceptance and favorable outcomes [5].

Other statistical models and techniques proven to be successful were not attempted in this study. In telecommunications, SVM was selected as the best model and claimed to be suitable for non-linear, non-normal, high volume and dimension of datasets with the main challenge of selecting the appropriate kernel functions and model parameters. This was attributed to SVMs simple classification plane, strong generation ability and good fitting precision [18] [9]. In [6], an Ant Colony Optimization technique (Ant-Miner+) and a rule extraction algorithm (ALBA) was used with SVM. Ant-Miner+ allowed domain knowledge to be included in the final rule set and ALBA used simple rule induction techniques and specific concepts of SVM.

Extra data mining techniques and algorithms were also reviewed. In [8], cost-sensitive learning was introduced by weighing of misclassified predictions and balancing of the minority class samples. In [19], a hybrid technique was used by combining two artificial neural networks (ANN). A back-propagation ANN filtered out unrepresentative training data with a self-organizing map ANN performing prediction modeling.

## V. Conclusion and future work

In this study, we developed a client churn prediction model using logistic regression, random forest and C5.0 decision tree with all models yielding prediction accuracies more than 90%. C5.0 was selected because of marginally better accuracy and model fit that can readily be adapted and implemented. The prediction accuracy of C5.0 can be improved further by using weighted misclassification cost for its boosted trees.

As stated, the aim is to develop an initial model that can be improved by expanding the attributes more applicable to home-based care services as described in the previous section. In addition, if the data quality issues in the case company are addressed, important attributes can be included in the modeling. With maturity on churn and retention models, integration into a companys enterprise system can be automated such as a churn amelioration system for a telecommunication company described in [15].

In a tournament on developing client churn systems participated by practitioners and academics described in [12], it was concluded that logistic regression and tree approaches performed well and are good techniques, to begin with by companies starting up a predictive modeling function. Exploring several estimation techniques to develop one model may not pay off.

## Acknowledgment

We would like to thank the anonymous Australian company for providing data and business domain expertise for this research project. And also for my colleagues Chumning Liu, Bin Fu and Xiao Zhu at UTS Advanced Analytics Institute for assistance in prediction modeling.

## References

[1] Davood Golmohammadi and Naeimeh Radnia, *Prediction modeling and pattern recognition for patient readmission*, International Journal on Production Economics, 171 (2016) 151-161.

[2] Forte and Rui Miguel, *Mastering Predictive Analytics in R*, 1st ed. Packt Publishing Ltd., Birmingham UK, June 2015.

[3] Y. Huang, F. Zhu, H. Yuan,et al., *Telco Churn Prediction with Big Data*, Soochow University, China, 2015

[4] H. Singh and H. V. Samalia, *A Business Intelligence Perspective for Churn Management*, Procedia - Social and Behavioral Sciences 109 (2014) 51-56.

[5] Kim Y. S., Moon S., *Measuring the success of retention management models built on churn probability, retention probability and expected yearly revenues*, Expert Systems with Applications, 39 (2012), 11718-11727

[6] W. Verbeke,D. Martens, C. Mues, B. Baesens, *Building comprehensible customer churn prediction models with advance rule induction techniques*, Expert Systems with Applications, 38 (2011), 2354-2364.

[7] Kristof Coussement and Dirk Van den Poel, *Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers*, Expert Systems with Applications, 36 (2009), 6127-6134.

[8] Xie Yaya, Xiu Li et al., *Customer churn prediction using improved balanced random forests*, Expert Systems with Applications,vol. 36 (2009) 5445-5449

[9] Kristof Coussement and Dirk Van den Poel, *Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques*, Expert Systems with Applications, 34 (2008) 313-327.

[10] Han J., Kamber M., *Data Mining Concepts and Techniques*, 2nd Edition, University of Illinois Urbana-Champaign, 2006

[11] S. Gupta, D. Hanssens, B. Hardie et al., *Modeling Customer Lifetime Value*, Journal of Service Research, 9 no. 2,(Nov 2006) 139-155.

[12] S. Neslin,S. Gupta, et al., *Defection Detection: Measuring and Understanding the Predictive accuracy of Customer Churn Models*, Journal of Marketing Research, XLIII (May 2006), 204-211

[13] S. M. Geron, K. Smith, et al., *The Home Care Satisfaction Measure: a client-centered approach to Assessing the Satisfaction of Frail Older Adults with Home Care Services*, Journal of Gerontology, 55B, no. 5 (2000), 259-270.

[14] Breiman, L. (2001), *Random Forest.Machine Learning*, 45(1), 5-32.

[15] McCausland et al., *Patent Number 5822410 - Churn Amelioration System and Method*, United States Patent, 1998.

[16] Van Rijsbergen, C. J. *Information Retrieval* 2nd Edition,Butterworth (1979).

[17] Anne Ruiz-Gazen and Nathalie Villa, *Storm Prediction: Logistic Regression vs Random Forest for Unbalanced Data*, Institut de Mathmatiques de Toulouse and Gremaq, Universit Toulouse, France.

[18] Xia Guo-en and Jin Wei-dong, *Model of Customer Churn Prediction on Support Vector Machine*, Systems Engineering  Theory and Practice, 28 issue 1,(January 2008) 71-77.

[19] Chih-Fong Tsai,Yu-Hsin Lu, *Customer churn prediction by hybrid neural networks* Expert Systems with Applications 36 (2009) 1254712553.

[20] Chih-Fong Tsai and Mao-Yuan Chen, *Variable selection by association rules for customer churn prediction of multimedia on demand*, Expert Systems with Applications 37 (2010) 20062015.

[21] P. Jermyn, M. Dixon and B. Read, *Preparing Clean Views of Data for Data Mining*, London Guildhall University and CLRC Rutherford Appleton Laboratory, UK.

[22] Y. s. Kim, H. Lee and J. Johnson, *Churn management optimization with controllable marketing variables and associated management costs*, Expert Systems with Applications 40 (2013) 21982201.

[23] Refaeilzadeh P., Tang L., Liu H., *Cross-Validation*, Arizona State University, USA.

[24] Norman Matloff, *The Art of R Programming*, No Starch Press, San Francisco, CA, USA, 2011.