POLITECNICO MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE

EXECUTIVE SUMMARY OF THE THESIS

Assessing partial defection of retail consumers, and the role of private label in its prevention.

TESI MAGISTRALE IN MANAGEMENT ENGINEERING – INGEGNERIA GESTIONALE

AUTHORS: Lorenzo Perego, Leoni Mirko

SUPERVISOR: Lucio Lamberti

CO-SUPERVISOR: Emanuele Fedrigolli

ACADEMIC YEAR: 2021-2022

1 Introduction

The discipline of customer relationship management involves, in one of its facets, the search for elements and models that will better enable the understanding of the phenomenon of churn/partial defection. This phenomenon can be identified with the final stage of the relationship between a customer and a given company. "Churn," in fact, refers to the instant when the customer stops purchasing the product or service from the company under analysis. In truth, the concept of churn takes on different declinations depending on the type of market in which the transaction takes place. Two types of purchase modes can be distinguished: purchase in which the customer-firm relationship is defined in contractual terms (known in the literature as "contractual settings"), and purchase in which the customer-firm relationship is not defined at the pre-purchase stage through the conclusion of a contract (known in the literature as "non-contractual settings"). The key difference between these two types of purchasing lies in the fact that in contractual settings the company has full visibility into the customer's churn, this is because the customer itself must formally withdraw from the contract to terminate the supply relationship (e.g., telco industry). In contrast, buying in non-contractual settings leaves the company with no direct signal about the customer's churn: the customer has no type of contract from which it must sever, the customer itself therefore does not expressly issue any signals confirming the churn. One type of industry in which non-contractual setting applies is that of fast-moving consumer goods (FMCG), the model disclosed here is applied in this contest.

For the reason given above, under non-contractual settings it results of chronic difficulty to stick to the traditional definition of churn, hence it is introduced the concept of "partial defection." Partial defection is alerted when the client under analysis demonstrates a significant reduction in its purchasing habits and enters a state of potential churn.

The churn/partial defection issue is of paramount importance to the business: being able to maintain a healthy relationship with the customer over time ensures prolonged cash flow from the customer. "Over time" is what aligns marketing strategy to firm's value: firm's value is created by present and future cash flows, which are eventually generated by each customer in the firm's customer base, those cash flows must be sustained by consistent marketing actions able to lock-in customers and boost the quality of their relationship.

A better understanding of the churn/partial defection phenomenon, from a managerial point of view, implies a greater ability to implement effective marketing actions aimed at preventing this phenomenon: being able to increase customer retention means increasing the return on the investment made in customer acquisition; each customer acquisition is in fact associated with an acquisition cost, which for the firm represents the investment made in the customer relationship.

In addition, the use of a model able to classify the customer base into risk classes, with relative probability of membership, can make a significant improvement in the addressability of marketing actions and thus increasing the yield on marketing efforts.

The most popular method used for segmentation into risk classes is the Recency-Frequency-Monetary analysis. This is usually done imposing thresholds on some specific variables used to describe customer behaviour, typically Monetary and Frequency. In addition, the obtained classification is deterministic, in the sense that membership in each class is not expressed such as the probability of belonging to that class.

More in general, the models currently used in churn/partial defection require the classification into risk classes to be provided as input to the model itself: these, in fact, are unable to process directly from the data the necessary classification, which will therefore have to be produced a priori. This is not a problem in the case of contractual settings, in which it is possible to assign a clear flag to churned customer. On the other hand, in non-contractual settings, the non-information on the customer churn obliges practitioners to decide on how to classify customer.

The solution presented in this paper is able to produce a probabilistic classification of the customer base directly from customer transactional data. Once the model has found the states of the Hidden Markov Chain (HMC), which correspond to the classes of risk, it is able to compute the transition probabilities between the identified classes. Furthermore, these transition probabilities can be combined with covariates via a multinomial logit model; this allows the overall model to take into account variables that may help in assigning the probabilities of customer membership to classes. In addition, by analysing the coefficients of the multinomial logit model, it is possible to understand what impact each covariate has on the probability of transitions between classes. In particular, by evaluating the impact of these covariates on the probability of transitioning from a higher class to a lower one (the latter with higher churn risk), it is possible to identify churn predictors. In this paper, using the procedure just described, it was found out that the covariates regarding the percentage of private label products bought proved effective churn predictors and as a signal of customer loyalty. Moreover, the issue regarding the loyalty-enhancing potential of private label products is still an open topic of discussion in the literature.

Furthermore, the solution proposed in this paper will be built, and successively tested, using real transactional data provided by an Italian retailer.

2 Literature review

A literature review has been carried out, in order to assess the current state of the art in the field of churn/partial defection modelling and prevention. The research looked for advancement in the statistical approaches and new findings in the field of churn prediction features, with special regard to the impact on brand loyalty generated by the conscious purchase of private label products.

The literature review was divided into three macro areas of interest:

i) Review on churn/partial defection modelling methods: this part reviewed the most widely used statistical models for studying the phenomenon of churn/partial defection. This review found that the models discussed in the literature are mainly applicable to contractual settings, where churner clients can be uniquely assigned with a label. These tools are primarily supervised learning models. In contrast, for non-contractual settings there is a dearth of models that can analyse the issue of churn/partial defection.

ii) Review on hidden Markov models: this review was conducted with the aim of studying the current and most innovative applications of hidden Markov models, even in areas not necessarily inherent to FMCG sector. This study allowed us to increase our knowledge of this innovative type of models, which was then exploited to build the model proposed in this paper. Furthermore, from the literature review it was determined that there is indeed a lack of application of hidden Markov models in the area of churn/partial defection analysis.

iii) Review on the role of private label on customer loyalty: the literature review of the effect that private label purchases have on customer loyalty was done with the aim of unearthing possible predictors to incorporate into our model. Actually, what emerged is that this issue is still open among researchers, and there are conflicting opinions regarding the loyalty-building power of private labels.

3 Research questions

From the literature review it was possible to came up with five research questions, that our disclosure will answer:

i) Develop a fully data-driven classification model that can assign each customer with a posterior probability of belonging to each of the identified classes. The model should work without any a priori classification.

ii) Identify, using the model, churn predictors obtained from the reprocessing of the information available in the transactional dataset, exploiting variables such as type of purchased products.

iii) Understand whether private label products have a role in churn prevention and churn prediction.

iv) Compare our model performances with the most traditional and widely adopted RFM models.

v) Provide useful managerial implications derived from the implementation of our model.

4 Methodology

A hidden Markov model (HMM) was implemented with the aim of obtaining an innovative and more precise customer classification respect to those that are currently in use. At the same time, an attempt was made to understand the role that private label products play in customer loyalty.

The development of the model went through several steps, which are report below:

i) Data Analysis: the width of the time horizon in which data are available, the choice of variables and the level of aggregation (i.e., choice of the time bucket) can influence the result of model parameters, hence it is necessary to understand which choices are most appropriate.

ii) Model specification and tuning: in this step, the variables identified in the previous step were selected, and if needed transformed. Several iteration cycles were conducted in order to find the final model. Indeed, this step was an iterative process, in which was involved a lot of trial and learn. The model was then improved by the addition of covariates.

iii) Output analysis: once found a satisfactory formulation of the model, a thorough interpretation of the results was conducted in order to understand the value of the information obtained as output.

iv) Model validation: to evaluate the capabilities of the model, a different dataset with different characteristics and completely new observations was tested. This step was done in order to assess whether the results obtained are consistent and whether the model is able to obtain meaningful states/classes and confirm the covariates impact.

v) Result discussion and managerial implications: the last step is to interpret the findings within the managerial context. In addition, will be disclosed the answers given to the research questions and possible future improvements.

5 Data

The data available for the study were retrieved from one of the largest FMCG players in the Italian market. The data were collected between August 2020 and September 2022. The dataset contains all transactions made by a pool of customers during the indicated period. The starting aggregation is at the ticket line level (i.e., each row corresponds to a unique product type in each ticket). The following variables are available: Card number, Monetary, EAN code, quantity, ticket date, and product description (in literal form).

6 Model specification and data management

The goal of the model is to classify customers into segments representing the different levels of churn risk and to understand which purchasing behaviour act as a churn predictor. Model training was done using Monetary and Frequency as continuous emission variables of the HMM. The choice of these variables was aligned with those most widely used in the literature to describe customer behaviour in FMCG sector. Indeed Frequency (number of visits per period) and monetary (€ spend per period) are two effective indicators of relation with the customer (a loyal customer visit often the store and spend an above average amount of money. A "defected" customer never or rarely come to the store and spend less money). In addition, they are two of the three variables that are used to perform the RFM analysis (most used method for segmenting customers in risk classes). This will allow a straightforward comparison between HMC model and RFM model. To train the HMM, it was necessary to aggregate the data into customer-specific time series. The aggregation in periods was done by monthly base, the driver of this choice was the trade-off between:

- Necessity of not to having too long-time buckets that would cause the model to be "slow" in detecting suspicious cases. Churn is a rather rapid phenomenon; too wide time frame can cause the model to not be able to grasp it timely.

- Necessity to not introduce potential doubts between an effective customer churn and an absenteeism period, which may simply be the consequence of a specific buying behaviour.

- Necessity to not over increase the computational difficulty: from a computational point of view short time frames cause poor reading of customer behaviour by increasing the probability of observing periods in which customer never visited the store (i.e., Monetary and Frequency values = 0).

Data cleaning operations were performed to eliminate those noise data coming from non-loyal customers, as well as those behaviours that are not feasible for a typical customer (i.e., cashiers who swipe their loyalty card resulting as customers with very high frequencies). After the outlier detection process, new variables were developed, starting from the information available on the purchased items (feature engineering), those variables will be used later as covariates in the HMM model. This was done with the aim of improving the performances of the model and, at the same time, discovering variables eligible as churn predictors and customer retention enhancers. The focus has been on private label products, which were further disaggregated into two product lines: named in this paper as "pvl_retailer" and "pvl_premium". For each of them the percentage of products type in the category respect to the total product types in the time frame selected was calculated.

The way these covariates were engineered deliberately does not regard either the purchased quantities either the total amount spent. Indeed, the objective is to capture the choice that the customer does when purchasing a private label product, notwithstanding the magnitude of that purchase which instead can be influence by external factors (for instance family size).

Eventually, the trained HMM has:

- 3 states representing high, medium, and low classes. Those correspond to low risk (Class 3), medium risk (Class 2), and high risk (Class 1) of churn/partial defection.

- 2 response variables: Monetary and Frequency through which states/classes are identified.

- 2 covariates (% of product bought from each private label line) that impact the probability of transition between states of the system. The meaning of these two covariates will be covered in the next paragraph 7.

7 Results disclosure

Impact of covariates on transition matrix

To improve the developed model, the use of covariates was introduced by means of a multinomial logit regression model built on transition probabilities. These covariates do not influence the creation of the HMC states, but only modify the transition probabilities between thwm.

The two covariates used, and their significance are listed below:

Percentage of "retailer private label products" (perc_pvl_retailer):

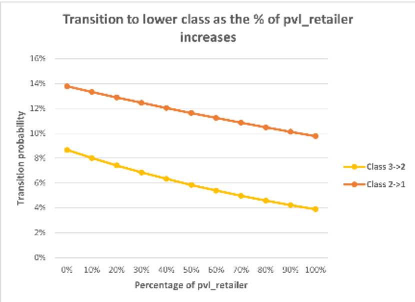
$$\frac{\#\ product\ types\ pvl_{retailer}}{tot.\#\ of\ product\ types}$$

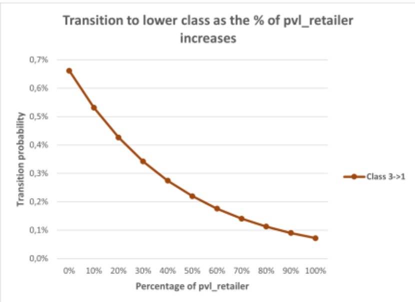
Percentage of "Premium private label products" (perc_pvl_premium):

$$\frac{\#\ product\ types\ pvl_{premium}}{tot.\#\ of\ product\ types}$$

For number of private label product types, it is intended the number of unique EANs referred to private label products, which is then divided by the total number of unique EANs in the ticket.

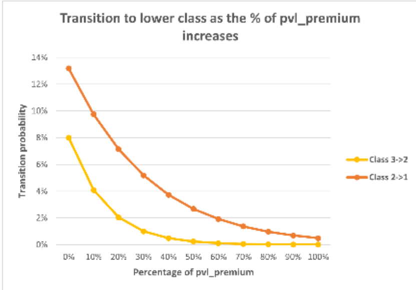
Effect of perc_pvl_retailer

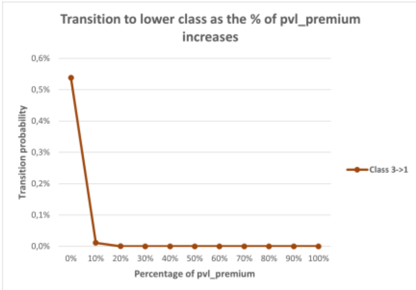




The covariate perc_pvl_retailer depends on the number of pvl_retailer products purchased; these products are sold by the retailer itself, which brands them with its logo. The pvl_retailer products are a substitute offering to well-known brand products that the retailer sales under his own brand. These products rely on value for money: low prices due to lower promotional and product management expenses, and good quality guaranteed by the retailer's ability in selecting suppliers, know-how derived from its core business. It can be seen from the graphs above that as the percentage of pvl_retailer increases, there is a corresponding reduction in the probability of transition to a lower status. This means that a customer who is more likely to purchase pvl_retailer products is indeed more loyal to the retailer and has a lower probability of increasing its risk of churn/partial defection.

Effects of perc_pvl_premium:





The perc_pvl_premium covariate represents the percentage of a given customer's purchase in products of the highest-end line offered by the retailer. These products are characterized by high prices and uncompromising quality. A customer who is used to buy this type of private label products is likely to be a customer who expresses high trust in the retailer and is willing to pay a premium price to have the best product the retailer can offers. The choice of pvl_premium is a conscious one; the purchase is not made for the purpose of savings money (in addition to the higher-than-average price, these products are rarely discounted); in fact, the customer pays more because he or she recognizes the quality of the product and knows that it meets his or her needs.

Among the two covariates so far implemented, this is the one that demonstrates the greatest discriminating power; indeed, the graphs above shows that even a small increase in the percentage of pvl_premium leads to a substantial reduction in the probability of transition to the lower state. This means that a customer inclined to purchase pvl_premium will rarely be a customer with little loyalty to the retailer.

In conclusion, private label retailers and private label premium, can become extremely useful tools in the hands of the marketing manager, who can use them to create marketing campaigns designed to build customer loyalty and boost customer retention. At the same time a low-risk customer that reduces the purchase of pvl_retailer, or pvl_premium, may be facing a deterioration of its relationship with the retailer; from this point of view these two covariates act as a churn predictors.

Performance comparison with traditional models: HMM vs RFM model

RFM model	Mean Monetary	Monetary Variance
Class 3:	358.8156	207.7375
Class 2:	187.7554	128.9996
Class 1:	43.9602	57.0894
HM model	**Mean Monetary**	**Monetary Variance**
Class 3:	394.431	222.862
Class 2:	153.468	89.901
Class 1:	19.658	23.807

RFM model	Mean Frequency
Class 3:	13.156
Class 2:	5.313
Class 1:	1.436
HM model	**Mean Frequency**
Class 3:	13.654
Class 2:	4.702
Class 1:	0.837

The tables above compare the mean values of Monetary and Frequency for each of the three risk classes: the values for the HMM model are taken from the output of the model itself, since they correspond by construction to the values of the distributions on which the class/states are fitted. The values for the RFM model, on the other hand, are obtained by calculating mean and variance, over the available history, of all values for each class. The substantial difference in these two models emerges when comparing the mean values for the highest risk class: Class 1 has lower mean values of Monetary and Frequency in the HMM model than the correspondent in the RFM model (Mon = 19.7 vs 44.0, Freq = 0.84 vs 1.44). Having a Class 1 with high mean values of Monetary and Frequency means that the model often includes in the lower-class customers that potentially have performance typical of clients with lower churn/partial defection risk (i.e., not at risk). Eventually, from the comparison of mean values it confirms that the model with HMC succeeds in creating classes that are consistent and coherent with the models currently in use among practitioners, while still being able to precisely identify customers at high risk of churn/partial defection.

To continue the comparison between HMM and RFM, the transition matrices obtained from the two models were compared. The values contained in these matrices represent the transition probabilities between the individuated classes. In the case of the model with HMC such matrix is obtained as part of the model itself. For the RFM-based model, on the other hand, the transition matrix was calculated by considering for each period the transitions between the three classes.

TP Matrix HMM	Class 1:	Class 2:	Class 3:
Class 1:	80.50%	19.38%	0.12%
Class 2:	13.27%	83.74%	2.99%
Class 3:	0.52%	7.94%	91.55%

TP Matrix RFM	Class 1:	Class 2:	Class 3:
Class 1:	77.56%	19.80%	2.65%
Class 2:	35.58%	47.24%	17.18%
Class 3:	5.99%	20.57%	73.44%

Looking at the transition matrix of the HMM, it is noticeable that the diagonal (representing the probability of remaining in the same) has always higher probabilities than the respective ones in the transition matrix of the RFM model. This means that the classification done by the HMM turns out to be more consistent. Indeed, this classification is obtained by assigning the probability that the observed emission of Mon and Freq belongs to the fitted distributions. Compared to the RFM model, the HMM is able to eliminate those noisy transitions that are the result of exceeding, even by a negligible delta, the predetermined thresholds, that eventually are even the same for all customers in the given period. In the RFM model, moreover, it can be observed that Class 2 turns out to be a highly unstable class; the probability of staying in fact corresponds to 47.2% while the probability of changing class (i.e., moving to Class 1 or Class 3) is 52.8%, thus making it more likely to change class than to stay in it. Finally, in both models, Class 3, with the customers at the lowest churn/partial defection risk, is reasonably the one "most closed" to movements to and from Class 1.

Model reaction to topical customer behaviour

Below are shown four typical behaviours that a retailer's customers may engage in. The results produced by our model for each of these customers chosen as archetypes will then be discussed.

-Stable customer: during the observation period these clients have a constant behaviour, always staying in the same class or at most making few fluctuations between the highest and middle class (Class 3 - Class 2).

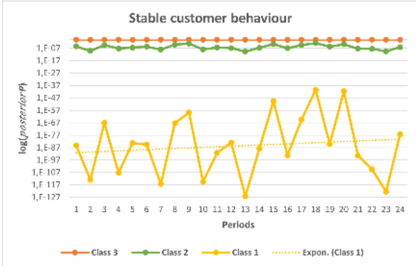
-Partial defection/announced churn: those are customers who show a gradual decline in their buying habits that leads them to enter the lowest state, Class 1. Those customers effectively go through a partial defection process, in which their performances decline period after period.

-Unexpected churn: for those clients the termination of the relationship is unpredictable. In fact, churn occurs quite suddenly and sometimes even after periods when the client's performance was even improving.

-Occasional customer: those customers over the observation time frames have highly variable behaviour. This archetype is indeed a very typical one for retailers: in the FMCG field it is common to observe a high level of variability especially if the time horizon of data aggregation is weekly or monthly (as in our study).

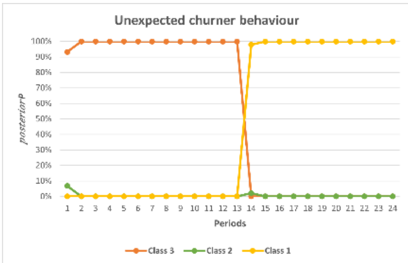
Below are reported the graphs showing how the model responds to the four archetypes. For comprehension’s sake a logarithmic scale graph will be provided. Indeed, logarithmic scale is useful to better understand the increase in the posterior probability of belonging to the lowest class. In fact, it is common for the value of the posterior probabilities of belonging to the lowest class to be extremely low (order of magnitude of E-100), especially when the customer is very loyal and thus belon to Class 3 with very high probability. Observing the increase in logarithmic scale allows to better grasp signals related to the change in the probability the customer has of transitioning from a higher to a lower class (with special regard to Class 1).

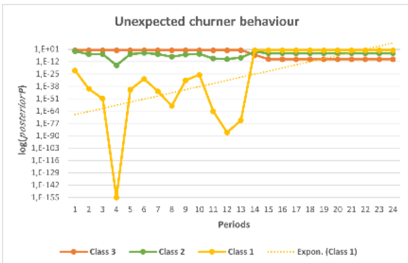
Stable customer:



In these graphs is shown the trend in the posterior probabilities that a stable customer has of belonging to each of the three classes. Specifically, the selected customer is assigned to Class 3, and remains stable in the same class for all periods. This means, as shown in the first graph, that the posterior probability of being in Class 3 is constant over time and has value very close to 1. Instead, in the second graph it is shown the evolution of the posterior probabilities of the same customer, but in logarithmic scale: the value of the posterior probability relative to the lowest class remains almost constant at very negligible values, the trend line is almost horizontal and is composed by extremely small probabilities.

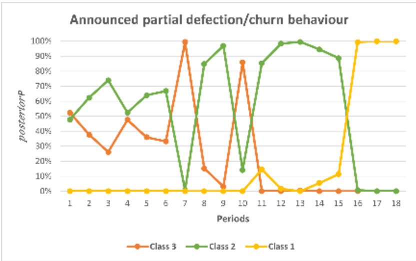
Unexpected churn:

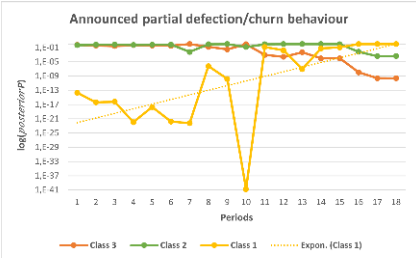




In this case the selected client, after a very stable period spent in the highest Class 3, suddenly churns. This behaviour is highly unpredictable and often caused by external factors, over which it is not possible to have any visibility or control. As can be clearly seen in the logarithmic graph, there are no signs that would point to a possible churn. This archetype is by far the most difficult to deal with; only the introduction of covariates to the model might help in recognizing this kind of sudden churns.

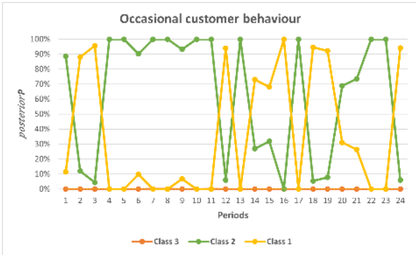
Partial defection:

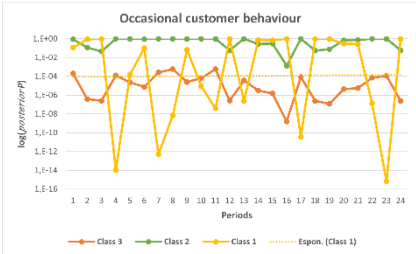




Customers who perform a partial defection are those clients that after some stable periods in one of the two high classes (or with small transitions between them) begin to show signs of relationship deterioration. In particular, in the first graph it is shown how the client during the first 10 periods was fairly constant, net of some acceptable variability. From period 11, on the other hand, there is a clear sign of the beginning of a partial defection process: the probability of belonging to the highest class is zeroed out and at the same time the probability of belonging to the lowest class increases significantly, even though the client is actually classified in the middle Class 2. This sudden increase is far more visible in the logarithmic scale graph: transition probability to Class 1 goes from orders of magnitude around E-20 and a peak of E-41, up to orders of magnitude ranging between E-01 and E-05. Instead, transition probabilities of belonging to Class 2 or Class 3 demonstrate a reverse trend (decreasing the probabilistic degree of membership to these two classes). The partial defection is even more evident observing the trend line of the probability of belonging to Class 1. Such trend line is pronounced upward, meaning that the associated risk of churn increases constantly during time. This customer represents the typical case in which the use of the model here disclosed would have allowed for an early warning of customer churn, that in the example from above are shown five months in advance.

Occasional customer:





Some customer behaviors are characterized by very frequent transitions between classes (often between Class 1 and Class 2). It can be seen from the posterior probabilities graph that the client in question is never assigned to Class 3 but rather continues to move from Class 2 to Class 1 and vice versa. In these cases, it is not easy to tell whether such customer is actually a churn/partial defection. However, it is still possible to carry out marketing actions aimed at stabilizing the customer in Class 2 and keep his behavior away from churn.

8 Model validation

In order to verify the stability of our model, additional transactional data were retrieved from the retailer’s management system; in particular, these represent new customers never used before. This step is needed to evaluate the model behaviour as the data provided change. A new hidden Markov model will be trained with the new data and compared with our master model.

The new dataset includes 6549 new customers. Like the main dataset this new one collects all the receipt lines purchased by new customers.

The main points to be checked are the following:

i) The model, with a fair variety of client, can recognize the three risk classes: high risk of churn/partial defection, medium risk and low risk. These classes are not likely to be numerically identical to those in the master model, indeed the values describing the classes obtained are strictly dependent on the input values on which the HMM is fitted.

ii) Consistency between the Monetary and Frequency values of the classes. The model is expected to be able to recognize that the class with lowest Monetary corresponds to the class with lowest Frequency, and so on.

ii) The lowest class must also be able to accommodate churner customers: the mean value of Monetary and the standard deviation of Monetary must be similar in absolute value in this way we are able to correctly classify even the cases of Monetary close to zero.

iii) The impact of covariates perc_pvl_retailer and perc_pvl_premium must be conserved. The new data should provide the same insights regarding the reduction of risk level as the percentage of private label bought increases. Moreover, perc_pvl_premium must have a higher impact than perc_pvl_retailer especially on the transition to Class 1, the one associated to highest risk of churn/partial defection.

Below are shown the classes of the test model

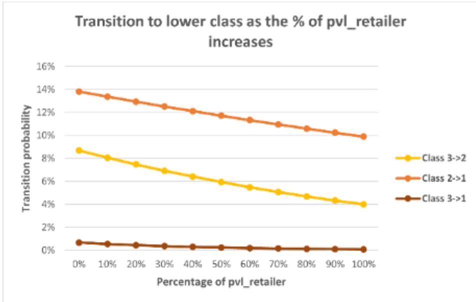
HMM test	Mean Monetary	Monetary Variance
Class 3:	246.078	144.208
Class 2:	130.675	80.738
Class 1:	26.342	23.584

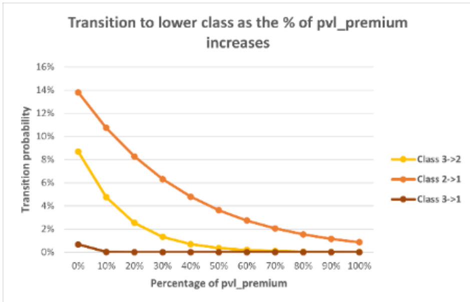
HMM test	Mean Frequency
Class 3:	16.5
Class 2:	7.0
Class 1:	2.3

It can be seen that the states found are consistent and coherent with the requirements listed above.

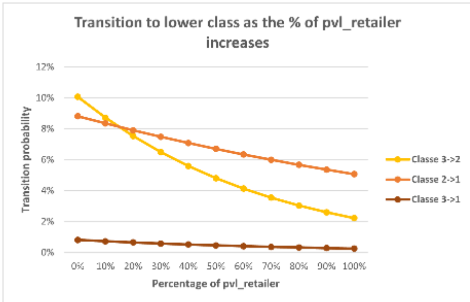
The trends of transition probabilities as the percentage of private label products increases are shown to be consistent and decreasing for both pvl_retailer and pvl_premium, with higher effect on pvl_premium. This means that the results obtained regarding the covariates are consistent and robust.

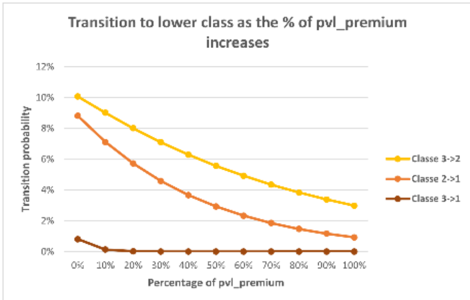
Master model:





Test model:





9 Conclusion, limitations and research boundaries

The model developed and discussed in this paper represents a potential starting point for subsequent improvements. Conceptually, the model may find further applications for other retailers in FMCG sector or in other non-contractual setting markets.

It should be emphasized that although the proposed model has excellent behaviour and resilience to "noisy" data, since it is still a data analysis tool the latter will perform all the better as the input data are cleaned of spurious data. In such terms, any improvement in the cleaning and outlier detection procedure could benefit the performance of the model.

Furthermore, as explained our model uses a normal distribution for fitting the Monetary of the three states of the HMM. From the knowledge acquired from working with retailer’s transactional data, the distribution of monetary in customers would appear to be best represented by a gamma-distribution. Such distribution could bring better consistency to the classes and make it easier for the model to create the aforementioned classes of risk. This because often the segments created, even with traditional models, presents distribution of monetary quite left skewed.

A final cue, from which future improvements can be implemented, concerns the integration and testing of new covariates such as product categories (i.e., frozen foods, fruit and vegetables, fresh meat, hygiene products). These, in addition to bringing a quantitative improvement to the classification capability of our model, can generate valuable managerial insights to be used as strategic marketing levers.

10 Managerial implications

The first managerial implication is due to the greater stability of the identified classes, which means being able to precisely divide the customers into classes that better describes their purchasing behaviour and the associated risk of partial defection/churn. In economic terms, this model can be used to identify a pool of customers which are more likely to churn and abandon the store. Consequently, it is possible to specifically targeting them with retention actions avoiding wasting time and effort on customers who do not present a high risk of defection.

The second managerial implication concerns the results obtained relative to private label products. Thanks to the evidence shown before, we can conclude that retailer private label products are able to induce customer loyalty. This statement is even more valid for premium private label products. With these findings, it is therefore possible to introduce private label products for loyalty marketing campaigns, trying to incentivize clients to purchase and taste these products. Moreover, the model itself provides an answer to the retailer about consumers' perceptions of private labels. In fact, it is reasonable to say that if retailer and premium private labels strengthen the relationship with the customer, it means that the customer himself appreciates them and finds in them adequate value for money. Thus, such products not only provide a higher margin to the retailer but also become a real marketing tool.

Another managerial insight arises from the covariate’s construction (which are calculated on monthly bases). This it is such that they are effective alarm bells for spotting the deterioration of customer relationship: if a good customer reduces its purchases in private labels, both retailer and premium, this can be noticed as an increase in the inherent risk of churn. Thanks to this implication the customer deterioration can be spotted in advance (i.e., churn prediction).

The last managerial implication concerns the validation of the model. On the one hand, it is easy to act by varying the type and the number of covariates in input to the model, accordingly to the result pursued for construction of marketing campaigns. On the other, as demonstrated by the model trained on new customers, the model has good flexibility to new data, and consequently resistance to the inherent variability in buying behavior, both in terms of monetary expenditure and frequency. In the end, the test with new customers validates the effectiveness of private labels in increasing customer retention; this shows that the result obtained from the Master model is for all intents and purposes attributable to actual buying behavior that discriminates high-performing customers from customers at high risk of churn/partial defection. From this result, it is possible to create ad hoc marketing campaigns to increase the degree of customer retention, particularly the customers in the high risk of churn class.

A final observation concerns the applications of the model: the model here disclosed was trained and tested on data from a single retailer, however, we have no evidence to affirm that the model is not applicable to data from other retailers active in FMCG sector: the metrics and KPIs used remain applicable, with appropriate adjustments, for any other retailer label. Moreover, the scientific-statistical basis of the model do not depend on the retailer from which the data was retrieved from.