AMERICAN MARKETING ASSOCIATION

2018 AMA Winter Academic Conference

Integrating Paradigms in a World Where Marketing Is Everywhere

February 23-25, 2018 | New Orleans, LA

PROCEEDINGS

Volume 29

Editors:
Jacob Goldenberg, IDC Herzliya
Juliano Laran, University of Miami
Andrew Stephen, Oxford University

2018 Winter AMA Conference

Integrating Paradigms in a World Where Marketing Is Everywhere

Volume 29

130 East Randolph Street, 22nd Floor • Chicago, Illinois 60601

AMERICAN MARKETING ASSOCIATION

Printed in the United States of America

Production Editor: Monica Gerhardt
Typesetter: Sarah Burkhart

ISBN 978-0-87757-367-8

Recency: Prediction with Smart Data

Florian M. Artinger, Max Planck Institute for Human Development
Nikita Kozodi, Max Planck Institute for Human Development
Florian Wangenheim, ETH Zurich
Gerd Gigerenzer, Max Planck Institute for Human Development

ABSTRACT
Since the early 1910s, managers have been using a simple recency-based decision strategy, the hiatus heuristic, to identify valuable customers. This study analyses the role of recency using a library of 60 data sets from business and other areas including weather, sports, and medicine. We find that the hiatus heuristic outperforms complex algorithms from machine learning, stochastic and econometric models in many of these environments. Moreover, if one includes further variables apart from recency in the complex algorithms, their performance does not improve. We show that the results are not so much driven by limited sample size than by the dominant role that recency plays in most of these environments. We conclude that less can be more, that is, relying on smart data such as recency can yield powerful predictions.

Keywords: *prediction, heuristics, behavior, experts*

Description: *We compare across 60 data sets a simple forecasting strategy that only relies on recency and has been used by marketers since 1912 against complex algorithms that make use of many variables. We find that the recency-based strategy performs exceptionally well.*

Predicting Future Events
Since the 1910s, marketing practitioners have used the recency-based hiatus heuristic for making predictions whether a customer will buy or not (Petrison, Blattberg, and Wang 1997). The notion of recency refers to a temporal metric which indicates the time passed since a specific action. The heuristic operates by comparing the time since the last event occurred to a fixed threshold, for instance whether a customer bought within the last six months. If this is the case, the heuristic makes the prediction that this event will also occur in the future. The hiatus heuristic is not bound to customers but can be applied to any domain where one makes a prediction whether or not an event occurs. For instance, it applies to a patient and whether she is readmitted to the hospital, whether a city will be hit by another tornado, or whether a sports team will win the championship again. In light of new developments in statistical modeling and machine learning, reliance on a single variable such as recency seems to be a strategy that easily can be improved upon. At the same time, there is a growing literature that documents that the simple heuristics that people have developed can perform well compared with complex methods given an uncertain environment (see for an overview Artinger et al. 2015).

This study analyzes the predictive performance of the recency-based hiatus heuristic across 60 data sets comparing it to state-of-the-art prediction methods. The purpose is not just to simply assert which strategy performs best overall but to shed light onto when it is commendable to use a complex method and integrate all available information respectively when it is best to rely on a heuristic and a single variable such as recency. It thereby addresses the important question when to use computerized methods or rely on human judgement of marketing practitioners and other experts.

Recency
The only information that the hiatus heuristic relies on is recency. Generally, a heuristic is a simple, psychologically plausible, decision strategy that ignores much of the available data. Just relying on recency can be regarded as an "extreme" heuristic, yet, reliance on recency is widespread in human judgement. Research on recency dates back to the 19th century when the "law of recency" was discovered which refers

For further information contact: Florian M. Artinger, Max Planck Institute for Human Development (florian.artinger@gmail.com).

to that memories of recent experiences come to mind more easily than memories from the distant past and are often the sole information that guide decisions (Brown 1838). People do not blindly apply recency but adapt their notion of recency to the structure of the task (Jones and Winston R. Sieck 2003). Yet, some authors suggest that the reliance on recency leads to maladaptive decisions. For instance, Kunreuther (1976) finds that people overreact to the occurrence of a natural disaster such as a tornado and attributes this to the reliance of recency. de Bruin shows that jury evaluations in Eurovision Song Contest depend on recency of the song in terms of its order of appearance (2005). Later performances receive significantly higher scores compared to the earlier ones, demonstrating positive serial position effects and challenging the fairness of the competition.

This negative picture contrasts with research in Marketing and the relevance of recency in prediction models in customer relationship management. The first use of recency were recorded in the 1910s in direct mailing when catalog companies developed the so-called "12-month prune rule": customers who did not make any purchase during the past year were dropped from the mailing list (Ross 1992). The rule is a specific instantiation of the hiatus heuristic that is still used in retail and banking sectors today (Miglautsch 2000; Wübben and von Wangenheim 2008; Persson and Ryals 2014). Wübben and Wangenheim (2008) show in three retail data sets that the hiatus heuristic performs as accurate or outperforms a modern stochastic model, specifically the Parteo/NBD model, in customer classification which rely on recency and frequency of purchases. Recency also plays a prominent role in machine learning algorithms for customer segmentation and CLV estimation. Neural networks, support vector machines, and hybrid models (Gupta et al. 2006; Tsai et al. 2013) typically rely on recency, frequency, and monetary value of the purchases integrating these with other variables for prediction.

In order to address when and why reliance on recency and the use of the hiatus heuristic can result in good judgement the literature on fast and frugal heuristics suggests two elements:

1. *Sample size:* the available data is relatively limited which implies that it is difficult to reliably determine the parameters of a complex model (DeMiguel, Garlappi, and Uppal 2007; Artinger and Gigerenzer 2016).
2. *Dominant variable:* a single variable dominates all other variables in terms of predictive performance and it is therefore sufficient to rely on this one variable irrespective of sample size (Martignon and Hoffrage 2002).

Data and Methodology

Data Sets

We gather a library of 60 data sets where each data set contains a time series of sequences of events (e.g., one sequence is whether or not a customer bought, whether or not a city was hit by a tornado) indicating for a given period whether an event occurred or did not occur. In the mean, a data set contains 2,268 (SD = 3,456) sequences of events, the mean number of time periods observed is 52 (SD = 36). Each data set contains between 3 and 10 variables that can be used to predict the occurrence of an event (mean = 5, SD 1.81). 24 out of 60 data sets are from a retail environment which allow us to verify the results from Wübben and Wangenheim (2008) and the performance of Parteo/NBD compared to the heuristic and to investigate the performance of other complex algorithms. 36 data sets are from non-retail environments such as the health domain (e.g., a patient being hospitalized, n = 4), weather (e.g., a country experiencing a sever drought, n = 2), sports (e.g., a team winning the NBA, n = 3), banking (e.g., conducting a money transfer, n = 5), crime (e.g., a country being hit by a terrorist attack, n = 4), and 18 other data sets (e.g., a country winning the Eurovision song contest).

Modeling Framework

Each data set is partitioned into two consecutive samples to form a calibration and a holdout sample. The calibration sample contains data on the first half of the time periods, while the holdout sample consists of the second half of the time periods. All sequences of events have one of two states: active if the event occurs at least once in the holdout sample and inactive otherwise. Each strategy makes a prediction whether an event will be active or inactive for the holdout period using the information in the calibration data. To train the strategies, the calibration sample is further divided into sub-calibration and sub-holdout samples which allow parameter calibration. The strategies compete in the holdout sample in terms of balanced accuracy which is defined as the average of correctly classified active events and correctly classified inactive events. It thereby accommodates type I and type II error. The framework is designed to simulate real-world forecasting experience. With sub-level data partitioning we only use information which is already available to predict future periods.

Prediction Methods

The set of prediction methods is constructed to include both simple and complex algorithms and allows us to analyze how strategies based on recency perform compared to other methods.

- *Hiatus heuristic*. Hiatus heuristic only uses recency and predicts that an event will be active if the event occurred last before a given time threshold
- *Random forest (RF)*. Random forest is an ensembling method based on a set of decision trees. It is one of the most widely used machine learning algorithms. Random forest does not require strict assumptions with respect to the data structure, which makes it a universal method applicable in many different areas.
- *Logistic regression (LR)*. Logistic regression is a well-established econometric method which is frequently applied to different classification problems in many environments including machine learning, social sciences, and medicine. It allows considering an arbitrary set of independent variables. We are using non-regularized logistic regression as this can be indicative of overfitting due to limited data.
- *Stochastic models*. Stochastic models that have been widely studied in the managerial and marketing literature. These models are designed to replicate the purchase behavior and assume that customers conduct their purchases at a certain stochastic rate. In our study, we use the two most widely used model modifications: Pareto/NBD and BG/NBD which both rely on recency and frequency as input variables.

The hiatus heuristic is a binary rule which requires a predefined recency threshold to classify the subjects. Due to limited space with focus in the following on results using a threshold value of T/2, where T is a number of time units in calibration sample. The remaining statistical algorithms estimate class-specific probabilities. Here, we classify subjects based on the natural 0.5 probability cutoff. We also evaluated the performance using the optimal threshold and the optimal probability cutoff, the results remain qualitatively the same but would require more space.

In order to analyze the importance of recency beyond its performance in the hiatus heuristic we consider three modifications of LR and RF models which differ by their predictor sets:

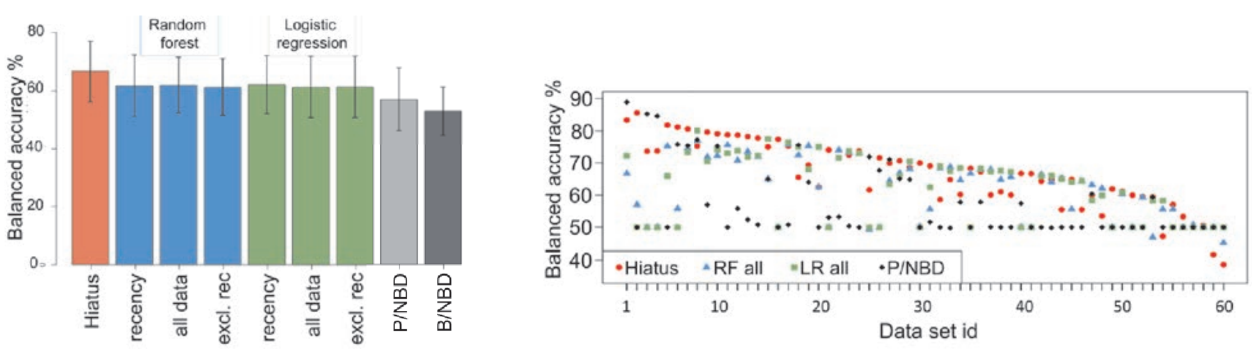
- *Model 1:* uses only recency as an input.
- *Model 2:* uses all available variables in a data set.
- *Model 3:* uses all variables except for recency.

Analysis

The performance of the different strategies can be inspected in Figure 1. Considering the left panel which shows the aggregate performance across all 60 data sets, one sees that the hiatus heuristic achieves the highest mean balanced accuracy. All complex algorithms perform worse, the difference is significant at the 5% level. The heuristic also has the smallest accuracy variance compared to all other algorithms except BG/NBD. The power of recency is also highlighted when considering the 3 variants of random forests, respectively the logistic regressions. Only using recency achieves on the aggregate level the same performance as when all variables are included. This suggests that there is little extra predictive power in adding more variables. At the same time, leaving away recency also results in about the same performance suggesting that the other variables can compensate when recency is dropped. On aggregate the results are virtually identical in retail and non-retail environments, hence we do not report them here.

The right panel of Figure 1 provides a more nuanced picture. It shows for each of the 6 data sets the performance of the hiatus heuristic, random forest and logistic regression when all variables are used and P/NBD. The Figure shows that

Figure 1. Accuracy Distribution



there is actually considerable variation between the different data sets and simply assuming that one should rely on the method that performs best in the aggregate would be misleading. The hiatus heuristic performs best only in 60% of the data sets. This highlights the importance to be able to point to when and why a simple strategy would be commendable compared to a more complex one that uses all the available information.

Table 1 shows the results of regressions on the performance difference between the hiatus heuristic and random forest/ logistic regression with all data, respectively Pareto/NBD. The table shows that neither sample size nor time units influence whether or not the heuristic performs better. Across the three different methods the main driver is whether or not additional variables apart from recency add any predictive power.

The dominant role of recency can also be demonstrated with variable importance measures. Variable importance refers to the random forest framework. It can be computed for each independent variable as a ranking measure which allows us to order the features by their predictive power. We have two variables which are contained in all data sets: recency and frequency, all other variables are specific to a given data set. Out of a maximum number of 10 variables, recency is the most important variable in 42 out of 60 data sets, it is the second most important variable in 10 data sets. In contrast, frequency is the most important variable in 10 data sets and the second most important in 39 data sets. Another reason why additional variables apart from recency do not improve predictive performance the random forest/logistic regression models is multicollinearity between the main variables. Mean correlation between recency and frequency is –0.52. Thus, once recency is already taken into account, further variables do not add much predictive power. On the other hand, exclusion of recency does not substantially decrease the accuracy, because the remaining variables are able to partly compensate the loss.

Discussion

This study analyzes the performance of the recency-based hiatus heuristic. We use a large data library to investigate how the heuristics performs compared to complex algorithms. Our results indicate that the hiatus heuristic can outperform complex prediction methods. This performance is not limited to the retail environment, where it was initially developed, but is observed in other environments as well. This is also the case for instance for predicting natural disasters such as tornados where research so far assumed that reliance on recency yields suboptimal performance (Kunreuther 1976). However, for instance in the prediction of the winner of the Eurovision Song Contest the information when a country won last and therefore recency has little predictive power. Instead the ordering of the contestants during the show is much more important.

The predictive power of the heuristic is not driven by small sample sizes but depends on the role of recency in the environment. The degree to which recency dominates in an environment boosts the accuracy of the heuristic and increases the gap between it and the complex algorithms. Recency proves to possess the largest predictive power in most of the data sets, facilitating good performance of the hiatus heuristic.

This paper touches on an important discussion: when should practitioners rely on computerized methods or their

Table 1. Regression on the Performance Gap: Heuristic Minus Complex Method

	Random Forest Gap	Logistic Regression Gap	P/NBD Gap
Sample size	.0001 (.93)	–.0004 (.28)	–.0001 (.86)
Time units	.00001 (1.00)	–.04 (.11)	–.003 (.86)
Share of recurring events	–3.49 (.82)	41.26** (.03)	–.73 (.97)
Share of recurring events squared	.80 (.95)	–40.60** (.02)	–4.18 (.75)
Ease of differentiation between events	.01 (.56)	.09** (.01)	.02 (.38)
Value added by additional variables besides recency	–.13* (.09)	–.23** (.02)	–.32*** (.001)
Observations	60	60	60
Adjusted R-square	0.01	0.28	0.23

Note: Table presents estimated coefficients and p-values in parentheses. * is $p < .1$, ** is $p < .05$, *** is $p < .001$

own human judgement. If practitioners can identify with reasonable confidence that in a given environment recency plays an important role they should rely on the hiatus heuristic. If this is not the case complex methods should be considered.

References

Artinger, Florian M., and Gerd Gigerenzer. 2016. "Heuristic Pricing in an Uncertain Market: Ecological and Constructivist Rationality." *Working Paper*, 1–47.

Artinger, Florian M., Malte Petersen, Gerd Gigerenzer, and Jürgen Weibler. 2015. "Heuristics as Adaptive Decision Strategies in Management." *Journal of Organizational Behavior* 36: S33–52. doi:10.1002/job.

Brown, Thomas. 1838. *Lectures on the Philosophy of the Human Mind*. William Tait.

Bruin, Wändi Bruine De. 2005. "Save the Last Dance for Me: Unwanted Serial Position Effects in Jury Evaluations." *Acta Psychologica* 118 (3): 245–60. http://www.sciencedirect.com/science/article/pii/S0001691804000885.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2007. "Optimal Versus Naive Diversification: How Inefficient Is the 1/N Portfolio Strategy?" *Review of Financial Studies* 22 (5): 1915–53. doi:10.1093/rfs/hhm075.

Gupta, Sunil, Dominique Hanssens, Bruce Hardie, Wiliam Kahn, V. Kumar, Nathaniel Lin, Nalini Ravishanker, and S. Sriram. 2006. "Modeling Customer Lifetime Value." *Journal of Service Research* 9 (2): 139–55. http://journals.sagepub.com/doi/abs/10.1177/1094670506293810.

Jones, Matt, and Winston R. Sieck. 2003. "Learning Myopia: An Adaptive Recency Effect in Category Learning." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29 (4): 626. http://psycnet.apa.org/journals/xlm/29/4/626/.

Kunreuther, Howard. 1976. "Limited Knowledge and Insurance Protection." *Public Policy* 24 (2): 227–61.

Martignon, Laura, and Ulrich Hoffrage. 2002. "Fast, Frugal, and Fit: Simple Heuristics for Paired Comparison." *Theory and Decision* 52 (1): 29–71.

Miglautsch, John R. 2000. "Thoughts on RFM Scoring." *Journal of Database Marketing & Customer Strategy* 8 (1): 67–72. http://link.springer.com/article/10.1057/palgrave.jdm.3240019.

Persson, Andreas, and Lynette Ryals. 2014. "Making Customer Relationship Decisions: Analytics v Rules of Thumb." *Journal of Business Research* 67 (8). Elsevier Inc.: 1725–32. doi:10.1016/j.jbusres.2014.02.019.

Petrison, Lisa A, Robert C Blattberg, and Paul Wang. 1997. "Database Marketing. Past, Present, and Future." *Journal of Direct Marketing* 11 (4): 109–25.

Ross, N. 1992. *A History Of Direct Marketin*. New York: Direct Marketing Association. Unpublished Paper.

Tsai, Chih-Fong, Ya-Han Hu, Chia-Sheng Hung, and Yu-Feng Hsu. 2013. "A Comparative Study of Hybrid Machine Learning Techniques for Customer Lifetime Value Prediction." *Kybernetes* 42 (3): 357–70. http://www.emeraldinsight.com/doi/pdf/10.1108/03684921311323626.

Wübben, Markus, and Florian von Wangenheim. 2008. "Instant Customer Base Analysis: Managerial Heuristics Often 'Get It Right.'" *Journal of Marketing* 72 (May): 82–93. doi:10.1509/jmkg.72.3.82.