

Paper SAS4301-2020

Optimizing Collection Strategies with SAS® Intelligent Decisioning

Luiz Kauffmann and Sunny Zhang, SAS Institute Inc.

ABSTRACT

Debt collection is an important part of asset resolution and credit loss mitigation for many institutions. Traditionally, collection decisions are based primarily on delinquency status, product type, and collateral type. In many cases, this can be a complicated and time-consuming process. As a result, recovery rates can be lower than expected due to lower collection revenue and high recovery costs. In this paper, we introduce a new collection decision process that identifies the optimal collection channel with the action using recovery action scores. The recovery scores estimate the “expected payoff” based on both customer-level risk models and channel propensity models incorporating cost variables. Micro-segments for customers are further developed to assist the mapping of specific actions to different customers. The approach is simple and intuitive, and we illustrate it with a case study applicable to both financial services and telco in the SAS® Intelligent Decisioning solution environment. This new action/channel scoring approach can help an institution design and optimize its collection strategies in order to improve recovery rates and collection revenue while reducing collection time and costs.

INTRODUCTION

The use of Collection Scoring models is a very common theme within credit risk management. Basically, any institution that has credit risk must necessarily create a process to manage its collection and asset recovery cycle. In the universe of companies that give credit to retail or telecommunications companies, the collection process needs to be automated. It also needs to present a good combination of efficient analytical models and a decision engine that has the ability to implement and automate the distribution of accounts and customers within the various channels. This analytical billing cycle needs to be constantly monitored and controlled. It needs to analyze each potential variable that can be used either as a new model input or the optimization of the collection rules. This continuous improvement activity ends up increasing the number of decision flows or even the number of segments over time. With more rules, the improvement process becomes more complex or often unfeasible to manage. Today, the financial industry has increasingly accepted the use of Machine Learning in their daily process. The main idea of this paper will be to present a practical proposal for the use of machine learning models combined with a simple and practical strategy to optimize collection flows. This strategy becomes easily implementable through the new SAS Intelligent Decisioning solution on the SAS® Viya Platform.

COMMON APPROACHES

In the financial industry, a very common practice is the development of collection scoring using a predictive model similar to credit scoring. Usually, this model is a logistic regression model in which the response variable is binary, indicating whether the defaulted account will be recovered or become a financial loss. This predictive model is usually transformed into a scorecard that classifies and sorts all the accounts or customers among different levels of

risk. For each level of risk (scoring bucket), one or two (champion and challenger) collection flows are created to assign a treatment or action to each delinquency level as determined by the “days past due” variables. For example, suppose you have a high-risk client as shown in the example below:

High-Risk Client:

- Days Past Due = 5
 - **Send a cell phone message**
- Days Past Due = 10
 - **Send a collection letter**
- Days Past Due = 20
 - **Send the customer to a call center channel that can guide the customer through the debt settlement process**
- Days Past Due = 30
 - **Send debt information to credit bureaus**
- Days Past Due = 45
 - **Send the customer to a specialized call center channel that can guide the customer through the debt settlement process**
- Days Past Due = 60
 - **Send the customer to a legal channel**

The most common approaches in the industry follow this same logic. However, it is very common to combine the risk levels with other variables such as type of product, geographic cluster, or type of collateral. As the natural improvement process continues, the emergence of new segments can cause an exponential growth in the number of collection rules. In addition, the performance measurement process will become more complex.

MICRO SEGMENTS AND ACTION SCORING

Within this process of adopting new analytical methods, new dimensions of modeling have been adopted through new concepts within the collection cycle. There are some models that are very common for the industry. Some common models are listed below:

Self-Cure

A predictive model that aims to forecast accounts or customers that do not need any aggressive collection action. These are of ten customers who have forgotten the due date of their debts. In some portfolios, this type of account claims to reach more than 70% of all debts with less than 30 days past due. This model is not necessary for scenarios where all accounts must have a self-payment configuration.

High-Risk Model

This predictive model aims to estimate the probability that a customer or account will become a financial loss given that the account is already overdue.

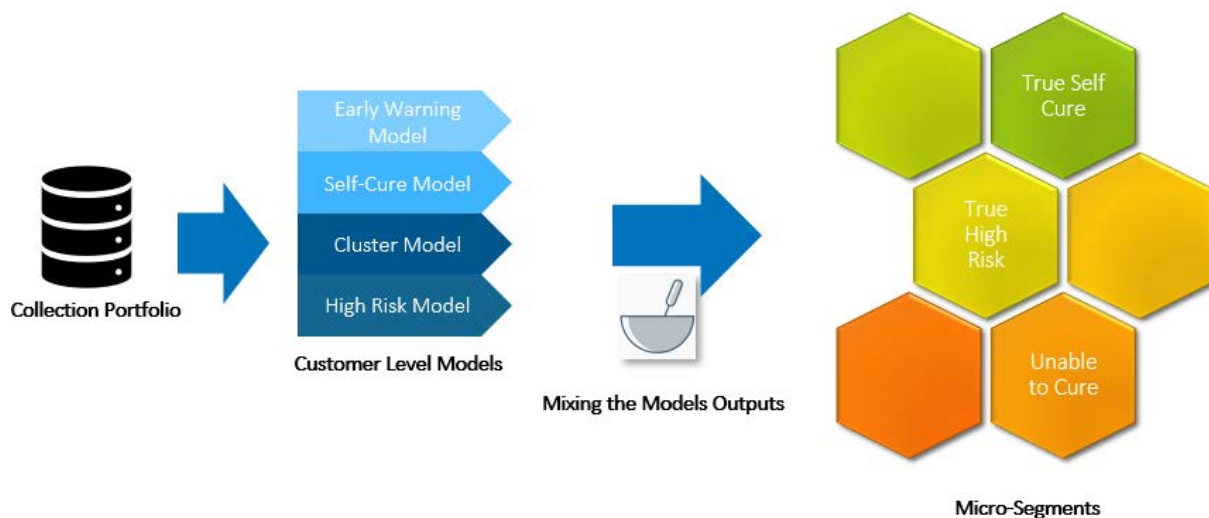
Segmentation Model

This model is basically the combination of an unsupervised model with a judgmental segmentation. It is very common to use techniques such as k-means so that different customer profiles are grouped in respective clusters. In addition, it is a common practice to isolate customers belonging to specific groups defined by the business area, such as FPD (first payment default) customers or other groups determined by the strategic area. After the unsupervised models are developed, exploratory analysis is recommended so that each cluster can be interpreted from the business perspective. For example, a cluster might consist of clients with an above-average income and multiple debts in the market.

Other models

Combined with these models, some companies also create models that capture other variables of interest in the collection process. These models might include Early Warning Models, Affordability Models, Best Time to Call Models, and others that depend on specific data and processes.

By combining the output of each predictive model and the segmentation model, we build what is called a micro-segmentation approach. The main idea of microsegments is the classification of customers into homogeneous groups considering their profile and their probability of recovery. With micro-segments, the collection action approaches can be parameterized according to the customer's micro-segment, and any optimization process can be performed more robustly. The micro-segmentation process can be illustrated as shown below:



The next step for the optimization of collection triggers is connected to the approach called Action Scoring, or in some places, Next Best Action. The central idea of Action Scoring is the development of models in which the target variable is the probability of recovering the account, given a specific collection action. For example, the model might determine the probability of conditional recovery when **the action is** "send a cell phone message".

For each treatment within the collection strategies, a predictive model should be developed. In addition to the propensity models, an expected payoff model can be considered for each treatment. These models will be necessary if each collection channel presents different options for debt settlement. For example, some channels may offer discounts to the

customer, so the payoff may have different distributions on different channels. This applies especially when the collection portfolio is distributed through external channels with different debt renegotiation policies.

The development of Action Scoring Models must be preceded by a strategic sample design. Below are the recommended steps for the sample design:

Step 1: Distribute the customers/accounts into each micro-segment.

Step 2: Conduct a randomized experiment where customers (inside of each Micro Segment) who have defaulted on their payments are randomly divided into N (number of actions used by the collection rules) groups:

- Customers who receive Phone call for recovery
- Customers who receive Email for recovery
- Customers who receive Direct mail for recovery
- Customers who receive "N-Action" for recovery

Step 3: Build two models for each of the segments of the customer:

- Probability to pay off/recovery
- Predicted pay-off dollars (only if the channels have different settlements or discount policy to recover the accounts)

Step 4: Score the entire customer base using the models built in previous step.

Step 5: Calculate the expected pay-off dollars given by the following formula:
 Expected pay off = (Probability of Recovery * predicted pay off dollars/given the channel) or
 (Probability of Recovery * Debt Amount/given the channel)

IMPLEMENTATION

After the development of micro-segments and Action Scoring Models, it will be possible to implement the collection strategy that can optimize the credit recovery process. To illustrate the logic, we have the following example:

Account id	Balance	recovery action: SMS		recovery action: Mail		recovery action: Bureau		decision
		expected pay-off	cost of recovery	expected pay-off	cost of recovery	expected pay-off	cost of recovery	
112	\$300	\$264	\$0.1	\$192	\$2	\$50	\$0.4	SMS
113	\$300	\$262	\$0.1	\$271	\$2	\$238	\$0.4	Mail
114	\$300	\$259	\$0.1	\$241	\$2	\$260	\$0.4	Bureau

expected pay-off = Probability of Recovery * predicted pay-off dollars / given the channel

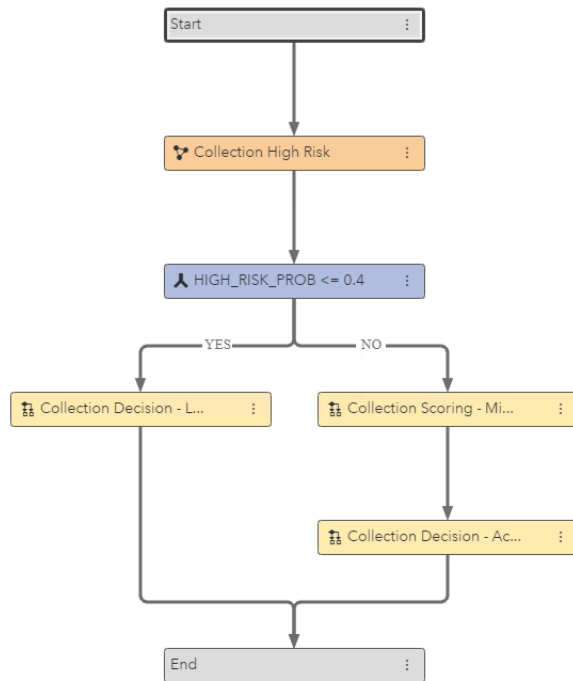
The implementation process will consist of a combination of four steps:

1. Assign customers to the micro-segments.
2. Within each segment, evaluate the client with all the Action Scoring Models.
3. Compare the expected payoff for each potential channel.

4. Assign the channel with the highest payoff as Next Best Action.

The implementation within SAS Intelligent Decisioning will be illustrated below:

Step 1:

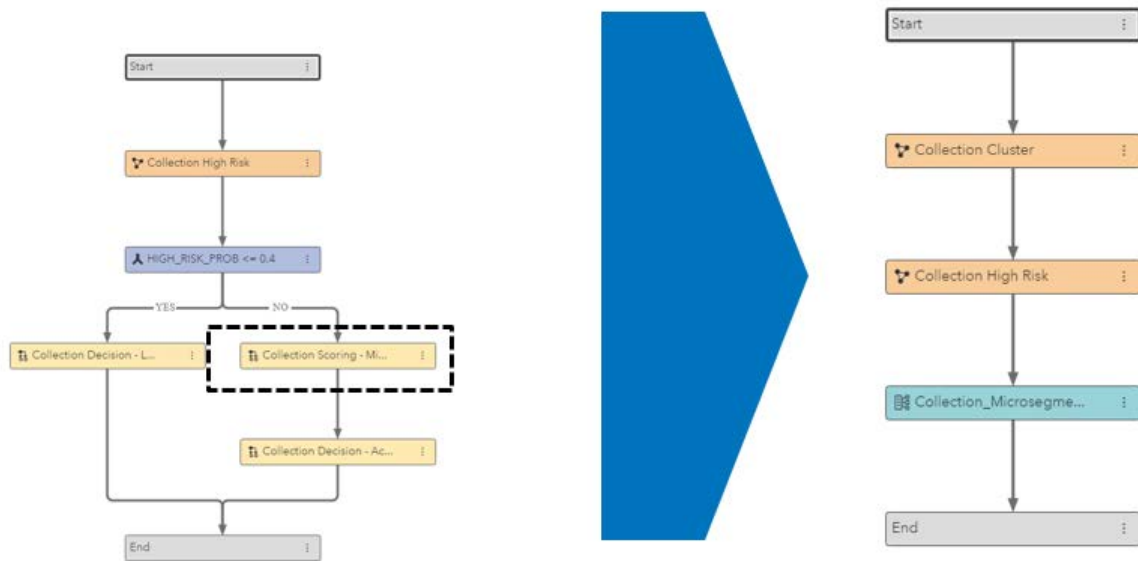


Decision flows are created visually, and the solution works with the object concept. Each node within the decision flow is an object. Each object can be one of following types:

- **Machine Learning model**
- **Logical test with one variable**
- **Call for other decision flows**
- **Set of assignment rules**
- **Query** of internal or external data sources

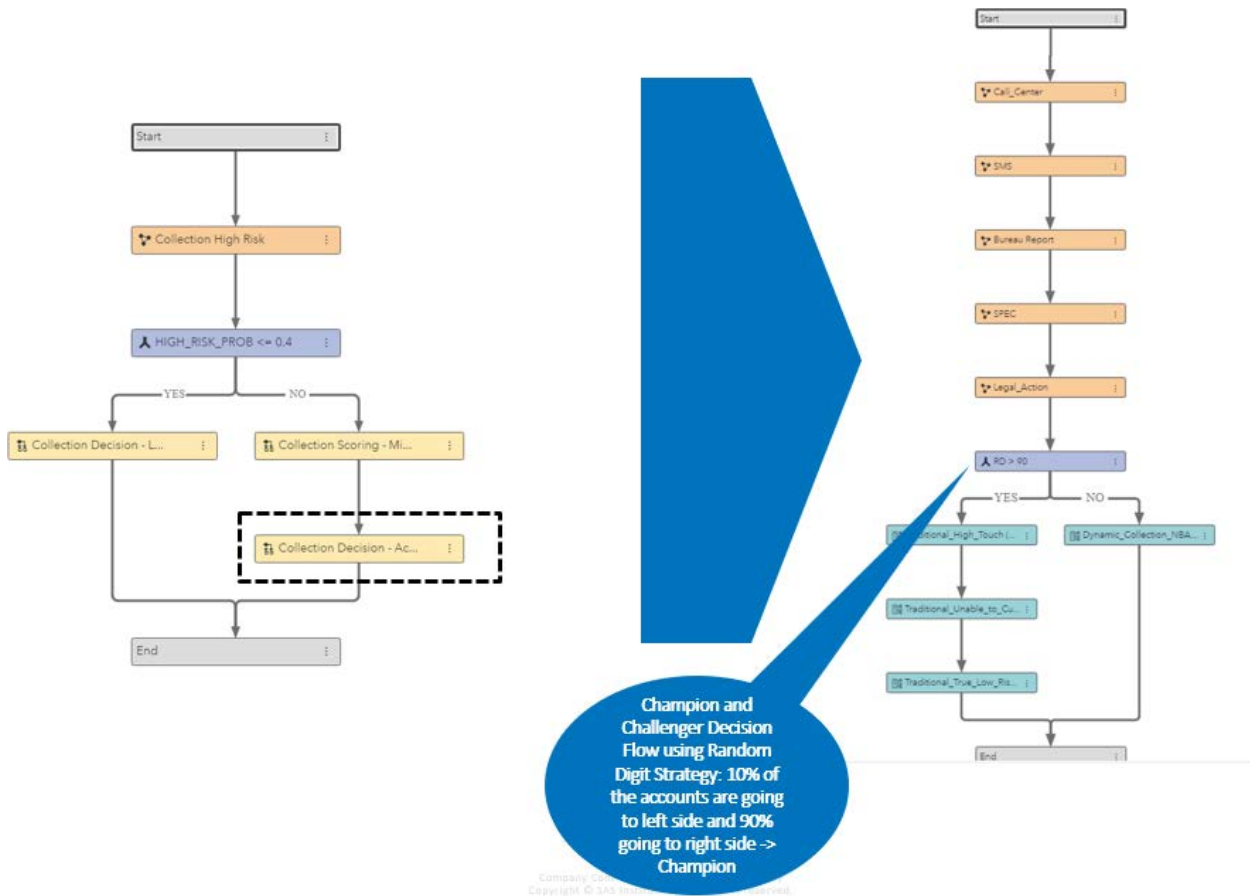
In the decision flow above, a High-Risk model is executed as a first step, followed by a branch in which customers with a probability score greater than 40% are directed to the decision subflows in the right path (the NO path).

Step 2:



The first subflow on the right side calls the segmentation model and the High-Risk model. These models create the micro-segments. The final rule set assigns each account to a micro-segment.

Step 3:



The second subflow in the NO path executes the Action Scoring Models, represented by the orange nodes. After the Action Scoring Models are executed, a random digit variable is assigned to each customer to split the customers/accounts into different strategies. In the current example, 90% of the customers will go to the Dynamic_Collection_NBA rule set and 10% (left side) will go to the traditional approach. The idea here is to represent a flow using a champion and challenger treatment.

Step 4:

The following figure shows a rule set that assigns the Next Best Action (NBA). It determines the NBA based on the days past due (DPD) level and by comparing the payoff between different channels.

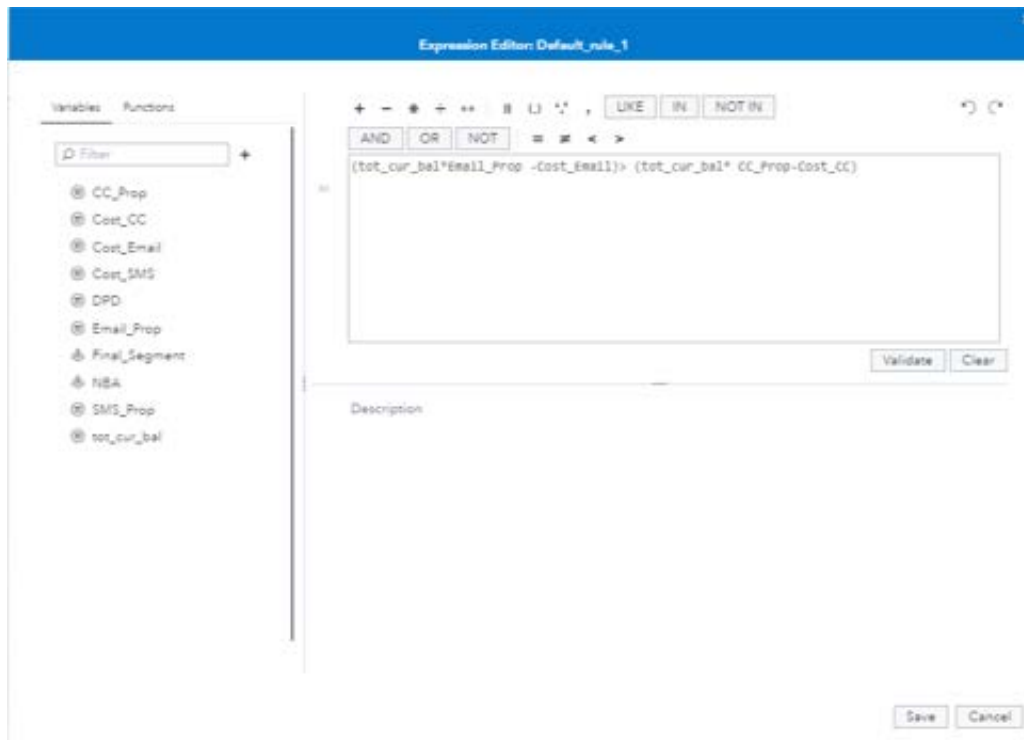
The screenshot shows the SAS Intelligent Decisioning - Manage Decisions interface for a rule set named "True_High_Risk (1.0)". The interface includes tabs for "Rule Set", "Properties", "Variables", "Scoring", "Versions", and "History".

The rule set contains four default rules:

- Default_rule_1:** IF DPD = 5 AND $(tot_cur_bal * Email_Prop_Cost_Email) > (tot_cur_bal * CC_Prop_Cost_CC)$ THEN ASSIGN NBA = "EMAIL".
- Default_rule_2:** IF DPD = 5 AND $(tot_cur_bal * Email_Prop_Cost_Email) < (tot_cur_bal * CC_Prop_Cost_CC)$ THEN ASSIGN NBA = "Call_Center".
- Default_rule_3:** IF DPD = 15 THEN ASSIGN NBA = "Call_Center".
- Default_rule_4:** IF DPD = 30 THEN ASSIGN NBA = "Call_Center".

At the bottom of the interface, there is a "+ Add Rule" button.

The user can edit more complex decision triggers such as this expected payoff calculation. In the following example, the expected payoff from the call center action is being compared to the expected payoff from the e-mail action. The NBA (Next Best Action) will be the one with the highest payoff, so an output variable called NBA is assigned for each account.



CONCLUSION

The approach presented in this paper requires some analytical maturity and a powerful decision engine. Because of this fact, many companies end up implementing only the micro-segment approach. The action scoring approach will require a big set of models for each action available and, depending on the decision engine, any new model can generate a big effort to translate the model logic. However, SAS Intelligent Decisioning is a new decision engine that requires a minimal effort to deploy any machine learning model and also includes some new decision capabilities. This collection optimization use case is a perfect example to illustrate this very practical use of SAS Intelligent Decisioning Solution. In addition, the platform allows the execution of other external models via APIs. The use case presented here can be treated as a first major step towards the collection optimization topic. It is also important to note that within the world of collection analytics most of these information sources carry a historical bias. Finally, this reinforces this use case as an introduction because all the Action Scoring Models are developed individually, so any potential bias can be analyzed.

Moving forward with the approach presented here, other advanced approaches can also be tested as a challenger strategy. There are some new studies to use Reinforcement Learning **methods as an alternative for "Next Best Action Models"**, and this will be a second step over the collection optimization topic.

ACKNOWLEDGMENTS

A special thank you to Gaurav Singh for his contributions to this use case discussion and knowledge sharing about the action scoring models.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Luiz Kauffmann
SAS Institute
Luiz.Kauffmann@sas.com

Sunny Zhang
SAS Institute
Sunny.Zhang@sas.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.