

Intelligent Collections: Order-to-cash powered by AI/ML



Abstract

Managing account receivables and prioritizing payment collections for B2B customers is critical for healthy order to cash and working capital cycles. Optimized strategies can directly impact critical KPIs such as : Revenue Leakage and Days Sales Outstanding (DSO). Traditional reactive strategies employed by businesses rely on greedy-optimization approaches that target large invoices after they go delinquent and do not pre-emptively collect at-risk future invoices.

This paper presents a framework that leverages enterprise big data, statistical models, and machine learning to enable intelligent payment collections. Specifically, we present the core techno-functional challenges and requirements and describe our framework that fulfills them to provide an acceptable, beneficial, and flexible solution for intelligent collections for B2B credit sales operations.

Introduction

An inefficient payment collection strategy negatively impacts working capital, causing liquidity problems, cash-flow inefficiency, toxic customer relationships, and manual effort overheads for collections. Since most organizations drive their sales through highly customized credit contracts, tracking and following up with customers is an effort-intensive process that is difficult to centralize and automate at scale. The entire order to cash process follows a cycle with every customer, which in totality affects the overall account receivables of the firm.

Order to cash process is primarily managed using standardized enterprise management solutions from vendors like SAP and ORACLE, which help capture and track invoices and their payments. However, such systems cannot implicitly build actionable intelligence for future payment cycles.

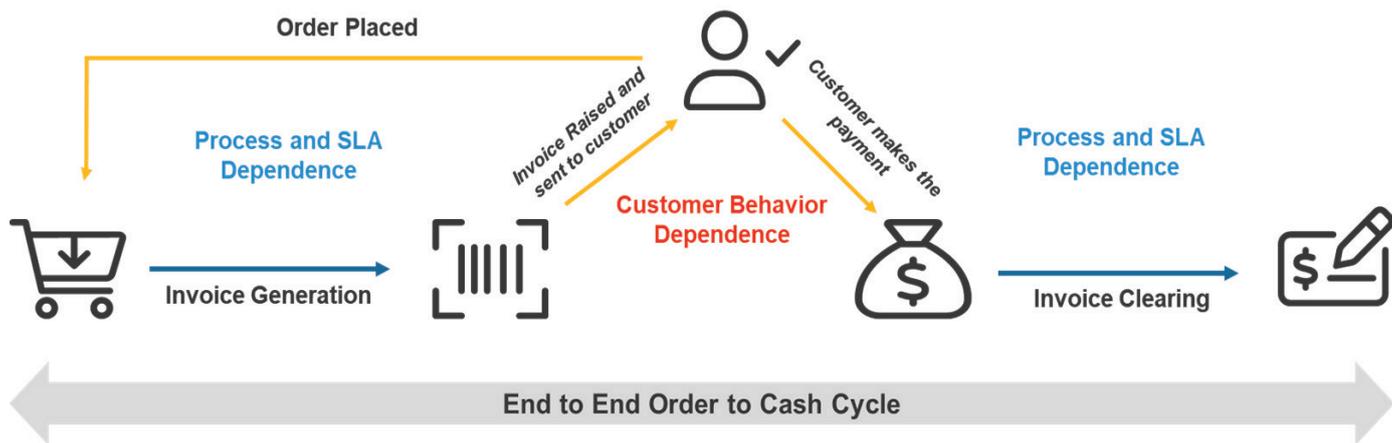


Figure 1: End to end order to cash cycle has a critical dependency on customers to pay on time

As shown in **Figure 1**, the key challenge involves failure to systematically predict and manage customer payment behaviour, preventing contract non-compliances, and increasing payment delays.

Including predictive intelligence into the collections requires businesses to focus on a few core KPIs and their associated indicators at the invoice or customer level, as shown in **Table 1**.

For each indicator, we can leverage data from OTC management systems and build data-driven predictions. In the following sections, we discuss the key challenges and requirements predictive intelligence needs to fulfill and elaborate on techniques used by our framework.

Key Indicators	Interpretation
DSO : Days Sales Outstanding	Average number of days to collect receivables. Measures cash flow health
WADTC : Weight Average Days to Collect	Average days between receiving an order and clearing the invoice. Measures payment efficiency of customer accounts
WADL : Weighted Average Days Late	Average days over term days that is taken by customer to pay. Measures delinquency in customers
Delinquency Rate	Percentage of invoices that are delayed/non-compliant for each customer. Measures overall contract compliance of the customer
Revenue Leakage	Interest revenue lost due to delayed payments. Measures the monetary impact of the customer account

Table 1: KPIs to consider to monitor health of OTC processes and collections

Our framework is standardized to work with any data lake/warehouse connected to enterprise data sources and is empirically tested for accuracy and efficiency with some of our partners. The illustrations and empirical evidence are based on experiments and outcomes achieved on actual data sets from our partners, which have been abstracted, anonymized, and re-sampled to test our framework while maintaining complete confidentiality.

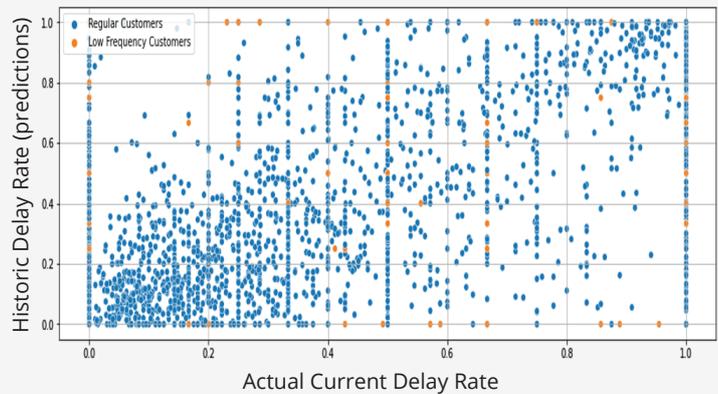
Critical requirements and challenges for predictive intelligence in OTC

Account receivables management systems capture transactional data, suitable for standardized reporting, but do not have features to derive intelligent KPIs to drive decisions. Any intelligence built for optimizing order to cash thus needs the capability to deal with these data gaps. Data and business process audits reveal several key requirements for developing a robust predictive intelligence solution.

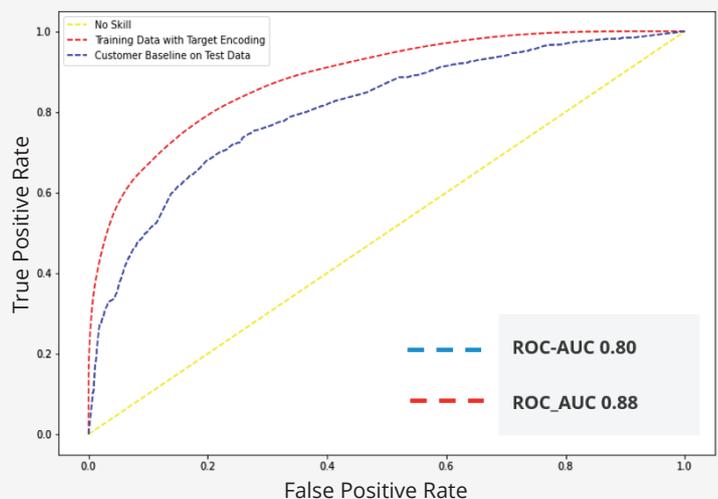
Need for isolating the signal from noise in account receivables data.

Careful data wrangling or additional external data is required to manage data correction for non-compliant customer behaviour. Examples include instances of delays due to weekends/public holidays, accounting process constraints, record errors. Such noisy data patterns trivialize risk categorization of customers using historical KPIs like WADL or delinquency rates, which in isolation fail to accurately predict future payments, as shown in **Figure 2**.

Figure 2: Using historical delinquency of customers as benchmarks produced a high variance and less accurate predictions



Historical delinquency and future delinquency has some positive correlation but variations are very high hence simple benchmarking techniques are not very accurate for future predictions



ROC-AUC Curve shows goodness of fit. Difference between red curve and blue shows that customers historical delinquency benchmark has a removable bias of ~8%

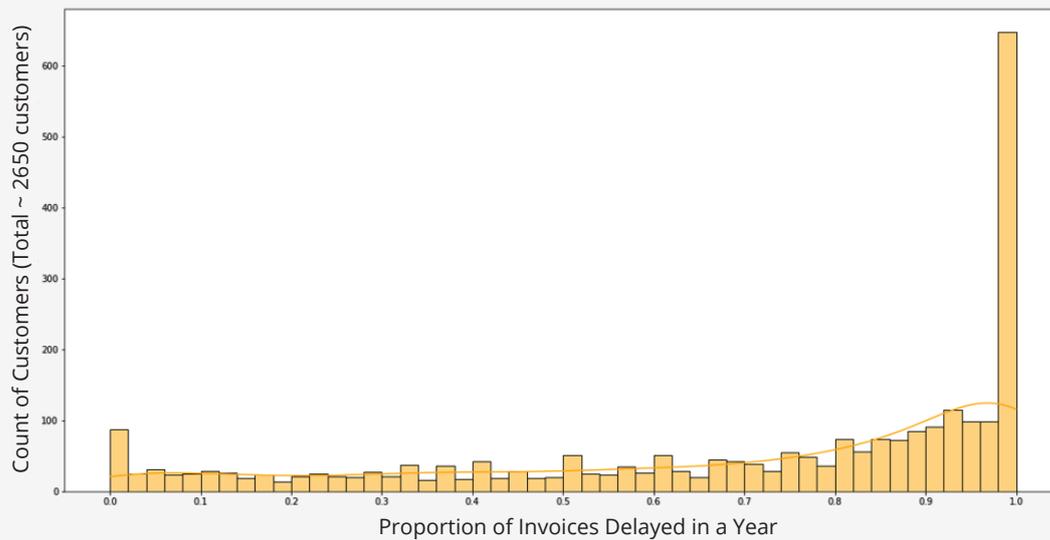


Figure 3: Customers delay (or not delay) consistently without interventions due to implicit behaviour/processes

Need of intelligent and stable indicators of customer behaviour.

Customers tend to consistently default (or not) because of implicit behaviours such as payment cycles and SLAs or evolving/pending disputes, as shown in **Figure 3**. Such factors are vital for predicting invoice delinquency and are not explicitly available in the transactional data sources. Strong causal features and predictors need to be mined out of the data to create accurate statistical or machine learning models.

Need for adapting predictive intelligence to paradoxical interventions.

Stakeholders consume intelligence on delinquency to intervene and alter customer behaviour (reduce future delays). Such a paradigm leads to concept and data drift over time, causing the performance of trained machine learning to degrade quickly. Therefore, the framework needs to continuously update the models with recent data to ensure stable and accurate predictions over time.

Need for adapting delinquency predictions to scenario changes.

Customer's risk is dynamic across the lifecycle of an invoice. Most machine learning approach use historical data to predict static estimates of delay/delinquency based on static inputs, i.e. invoice level predictions do not change over the tenure of the invoices. Successful tactics for collection need reliable predictions that consider the payment behaviour of customers as we move closer to the due date of the invoices.

Need for holistic risk prediction and aggregation at customer level.

Businesses mostly take holistic actions towards customers. It is necessary to aggregate risks at the customer level without losing granular intelligence at the invoice level. A holistic framework for collection optimization needs to deliver both strategic (payment policy) and tactical (collections) intelligence by scoring credit risks and predicting payment delays.

Framework for Predictive Intelligence

Our framework has several key modules for payment collections and customer delinquency management. The logical components of the framework are presented in **Figure 4**. We cover the salient aspects of each module in the following sub-sections

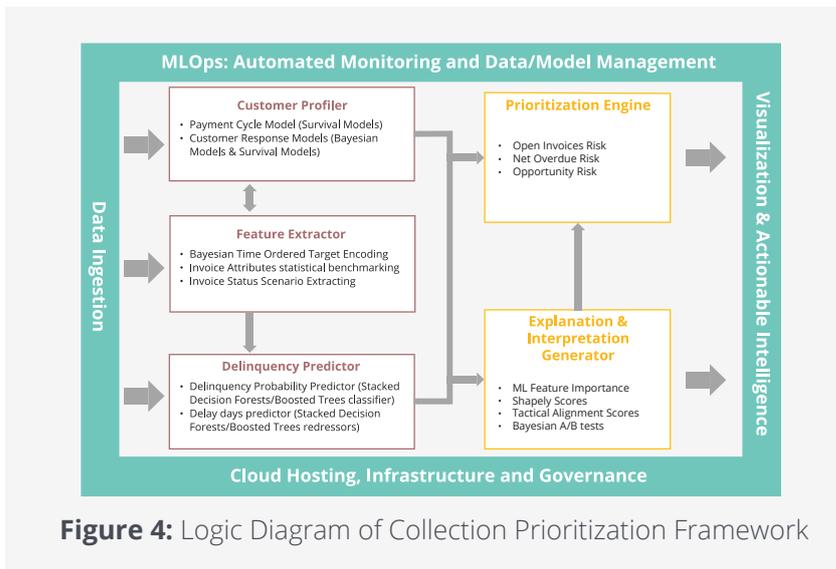


Figure 4: Logic Diagram of Collection Prioritization Framework

Feature Extractor

Account receivables data provide key invoice attributes such as contract terms, billing and due dates, invoice value, and payment dates. These raw data points, in isolation, have empirically shown low predictive power for payment delinquency, as illustrated in **Figure 5**. An extensive feature extraction engine is developed to mine powerful indicators using domain-specific and statistically validated hypotheses. The list of features the extractor generates is summarized in **Table 2** and provides a lift of ~20% over a basic Machine learning model. The feature extractor supports downstream machine learning and statistical models with variables that have high predictive power. Selection of historical time period for feature calculation typically varies between 3 to 6 months and is based on data and business process audits. The solution is operationalized to dynamically re-compute the features periodically based on the recent historical window.

Key Indicators	Interpretation
Customer Payment Profile	Average Payment Gap days, Most popular payment day of month, variance of payment cycles
Customer historical delinquency behaviour	Average delinquency rate and delay days, variance in delinquency rate and days of delay, average payment days
Customer historical invoice attributes	Average term days, average invoice amount, average invoices raised per month
Invoice creation scenario	Outstanding invoices during invoice creation, total overdue during invoice creation, amount due during the tenure
Categorial invoice attributes encodings	Proportion of delays by a customer on particular month/quarter/day, weight of evidence of channel/country
Invoice raw attributes	Amount, terms days, payment method, discount
Derived invoice benchmarks	Amount per day due, invoice amount to historical average terms days ratio to popular term days

Table 2: Summary of Features used for predictive modelling

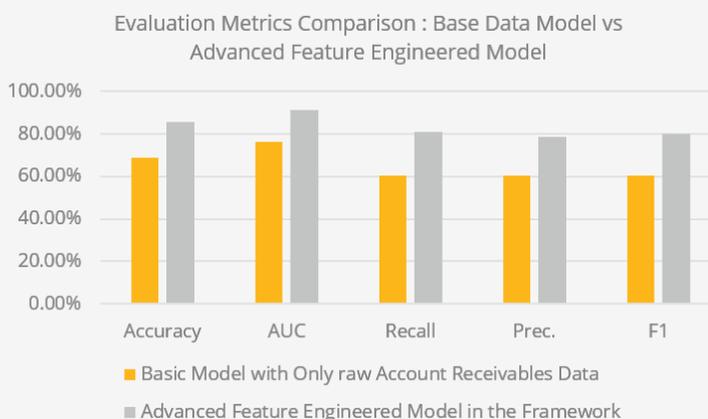


Figure 5: Lack of predictive power from basic account receivables data points

Our feature extractor utilizes state-of-the-art techniques for outlier correction and uses an iterative-imputer for missing value imputation. Customer features use time-ordered Bayesian target encoding and are dynamically re-computed (to capture the latest customer patterns). Bayesian encoding is an adaptation of state-of-the-art practices illustrated by catboost and mean encoding and delivers balanced machine learning models when dealing with a combination of regular customers (large sample of invoices) vis-à-vis new customers (a small sample of invoices). The presence of statistically powerful features helps the models generalize at higher accuracy levels with minimal overfitting and data leakage.

Customer Profiler

Account receivables systems capture customer identities, but not payment behaviour. Payment behaviour information is critical for - prioritization strategy since it is essential to connect with the customers at the right time for the right set of invoices and anticipate response times to avoid incessant follow-ups.

Our profiler uses a mix of Bayesian models and survival models (e.g., Kaplan Meier Curves/ AFT models) to quantify uncertainty in the payment behaviour of customers based on the most recent months of account receivables and collections workflow data. The key behavioral profile features extracted are summarized in **Table 3**. We use these models and derived metrics to categorize and compare customers based on payment cycles and risk across the collections life cycle, as illustrated in **Figures 6a and 6b**. These models help adjust delinquency and payment risks dynamically as invoice moves across its life cycle and feeds into the prioritization engine, enabling intelligent follow-ups with customers. In addition to statistically estimated parameters, the customer profiler calculates other functional KPIs from each customer’s raw data, such as outstanding credit deficit, disputed invoices, current net overdue amount, and ageing days, which further aid collection prioritization.

Key Indicators	Interpretation
Time to payment	Probability of payment (or not paying) on a specific day given his previous payment date
Time to respond	Probability of fulfilling particular response after follow up action has been taken by collection agents
Average Payment Cycle	Average time between 2 payments for the customer, most popular date of month for payment
Variation in Payment Cycles	Deviation from average time between payments and from observed popular payment days
Positive response Rates	Probability of paying by contractual dates given a certain action on the customer

Table 3: Statistical models and measures for customer payment profiles

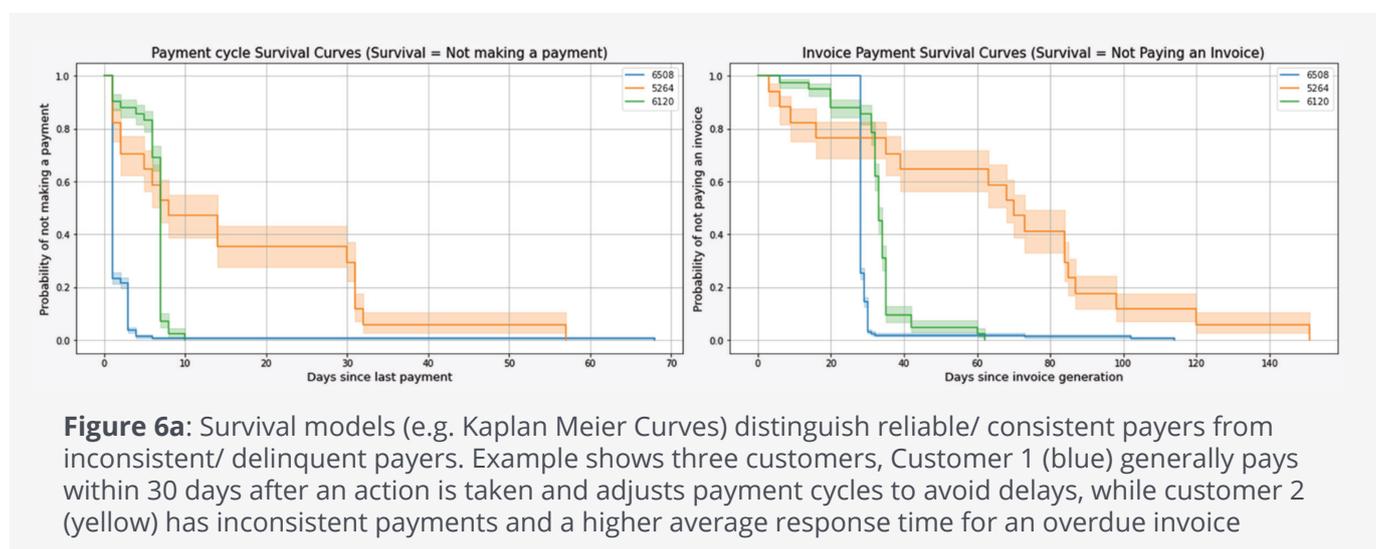




Figure 6b: Bayesian models can predict the response (e.g., promise to pay fulfillment, deduction claim requests) to tactics and help in daily downstream prioritization for collections. E.g., shows the probability distribution of response days for one customer across 3 sample tactics with typical days to respond as 7, 14, and 5 days with different variances, respectively.

Payment Delinquency Predictor

We use two machine learning models to predict outcomes on invoices. A classifier predicts the probability of an invoice to get delayed (non-compliant with term-days), and a regressor predicts the days of delay that can be expected, assuming the invoice gets delayed.

The models are fitted on a superset of approximately 100+ features, including major casual drivers of delayed invoices. The current design of the models uses an ensemble of decision forests and gradient boosted trees using a meta-learner (blended ensemble). Decision forests have shown the strong capability to deal with high-dimensional heterogeneous features with non-linearity and provide the best fit for the use case, as shown in **Figure 7**. We avoid Neural networks due to longer experimentation time, and higher compute resource requirements.

Using SMOTE and bootstrapping techniques, our models deal with class and categorical feature imbalances (e.g., payment methods).

Similar techniques are also used to reduce bias in the training data towards any category of customers. Models are retrained every month, and cross-validated using a time-ordered train-test split, ensuring the validation accuracy emulates the online inference scenario. Periodic online retraining provides adaptability to changes in the behaviour of delinquency due to strategic interventions.

Inferencing is daily or weekly for every open invoice with features updated with the latest data. These models help predict delinquency risk as soon as the invoice is created. We continuously monitor feature importance, distribution of the data, and model performance to manage and overcome performance stagnation.



Figure 7: Empirical evidence that ensembled Decision Forests/ Gradient Boosted Trees provide accuracy in the range of 83-86%

Prioritization Engine

The guiding principles of the prioritization framework are to deliver intelligence that can help customize invoice policies, maximize collection amount, minimize overdue risk, and optimize human effort for collections. Our prioritization framework thus periodically emulates the following questions:

Which customers need strategic interventions, and which need tactical attention?

Which invoices to be followed up on with every customer?

Which call-back efforts can be avoided or de-prioritized?

Our framework derives three metrics (illustrated in Figure 8) from covering current overdue, future risk, and customer value, considering time horizon and dynamic payment/collection factors.

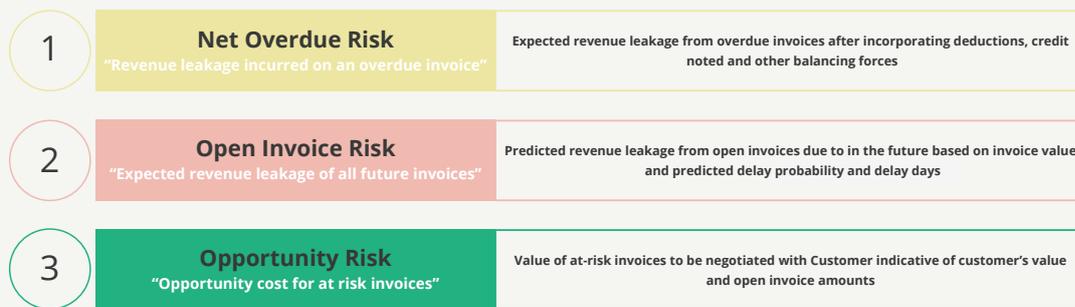


Figure 8: Three key metrics derived for prioritization. Metric-1 is calculated on actual current data, metric-2 uses predictions from the delinquency predictor (adjusted with customer profile models) metric-3 represents the value of the customer (invoices at risk)

Risks scores are computed at the invoice level and rolled up at the customer level to prioritize customers. Scores are derived using financially intuitive algorithms or business rules on top of the predictions from machine learning and statistical models. Delinquency model predictions are used to calculate open invoice risk in combination with invoice values, and net overdue risk is calculated from the current overdue amount and historical benchmarks of WADTC/AAD. These two risks are adjusted dynamically using payment delays, and response profile models, which provide additional probabilities for defaulting/delaying given the customer's updated payment cycle/behaviour. Customer value provides the third dimension, which is the opportunity size. **Figure 9** illustrates how these scores can practically drive intelligent collections.

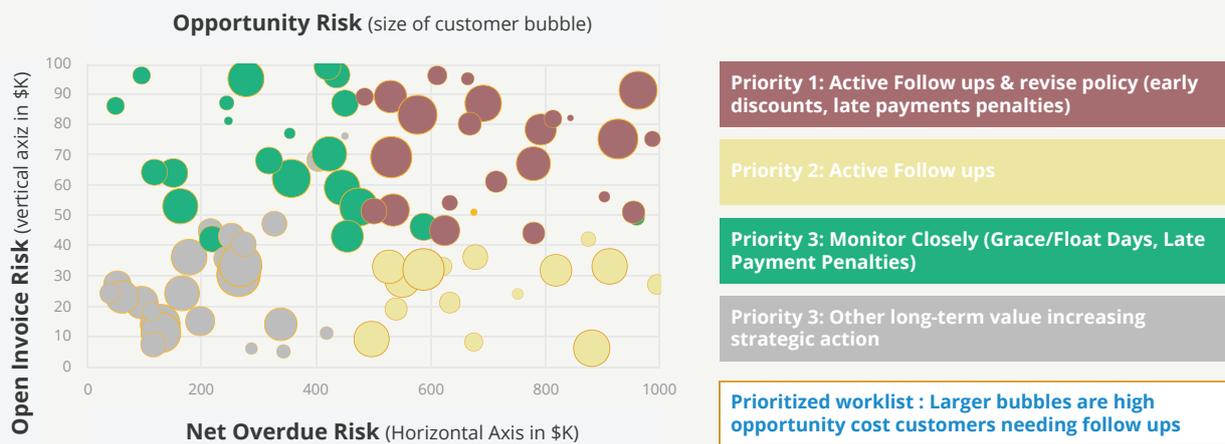


Figure 9: Illustration of serving risk scores to account receivables managers and collection agents for revising receivables policy and prioritizing customers for next best action. Additional dimensions for decision making can be easily embedded on top of these 3 key metrics allowing for customizations such as intelligence on customer type (shape/palette of the bubble), custom rules for omitting to follow up due to SLAs, or considering time since last intervention (shade of color).

Explanation and Interpretation Generator

The core engine of our framework is powered by machine learning and statistical models. Hence, predictions prima-facie are black-box in nature, especially for the end-user of the solution. To aid intuitive adoption, we encapsulate the following functionalities of explanation and interpretation into the solution:

Feature Importance scores for each machine learning model for guiding users on what features are essential and considered for the predictive functionalities.

Shapely Scores to intuitively explain invoice level predictions through the quantified influence of key factors.

Diagnostic KPIs such as trends on WADTC, DSO, Delinquency rate for customers to associate of historical performance and predicted KPIs.

Tactical alignment scores using statistical measures such as Kendall's Tau12 to indicate how prioritization is aligned (or different) from as-is or reactive strategies due to the introduction of predictive scenarios.

Bayesian A/B Tests for quantifying impact of using the intelligence framework over a specified duration.

Conclusion and Future Prospects

The framework is built on plug-and-play principles as several independent components integrate seamlessly to provide the necessary intelligence for decision-making. This enables an agile implementation and custom deployment roadmaps based on enterprise technology and data maturity. The framework can be deployed using open source and cloud compatible technologies like data lakes, Hive, SQL, Python, Spark, and React and is compatible with major cloud service or on-premises environments.

Current implementations have shown great potential in optimizing order to cash management and improving collections. The key areas for further research and development include:

Improving the efficacy of the models by incorporating additional data sources, e.g., CRM, customer emails.

Introduce feedback loops into our solution and create Bayesian bandits (reinforcement learning) to automate the next best action.

Building causal models to incorporate counter-factual predictions (i.e., what will happen if a particular action is not taken)

These advancements are complex and a part of the long-term roadmap for the solution and will further improve the accuracy and efficacy of the solution and create more business value.

Our Expert



Amartya sen

Senior Data Scientist,
AIML - ADM Fractal

Enable better decisions with Fractal

Fractal is one of the most prominent player in the Artificial Intelligence space. Fractal's mission is to power every human decision in the enterprise and bring AI, engineering and design to help the world's most admired Fortune 500® companies.

Fractal product companies include Qure.ai, Crux Intelligence, Theremin.ai, Eugenie.ai & Samya.ai.

Fractal has more than 2,300 employees across 16 global locations, including United States, UK, Ukraine, India, and Australia. Fractal has consistently been rated as the india's best company to work for, by The Great Place to Work® Institute, a 'Leader' by Forrester Research in its Wave™ on Specialized Insights Services, Computer Vision & Customer Analytics and as an "Honorable Vendor" in 2021 Magic Quadrant™ for data & analytics by Gartner.



Corporate Headquarters

Suite 76j,
One World Trade Center, New York,
NY 10007

[Get in touch](#)