FLOWCAST

WHITE PAPER
2018

BIG DATA, SMART CREDIT

CLOSING THE SME FINANCE GAP THROUGH ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

TABLE OF CONTENTS

EXECUTIVE SUMMARY

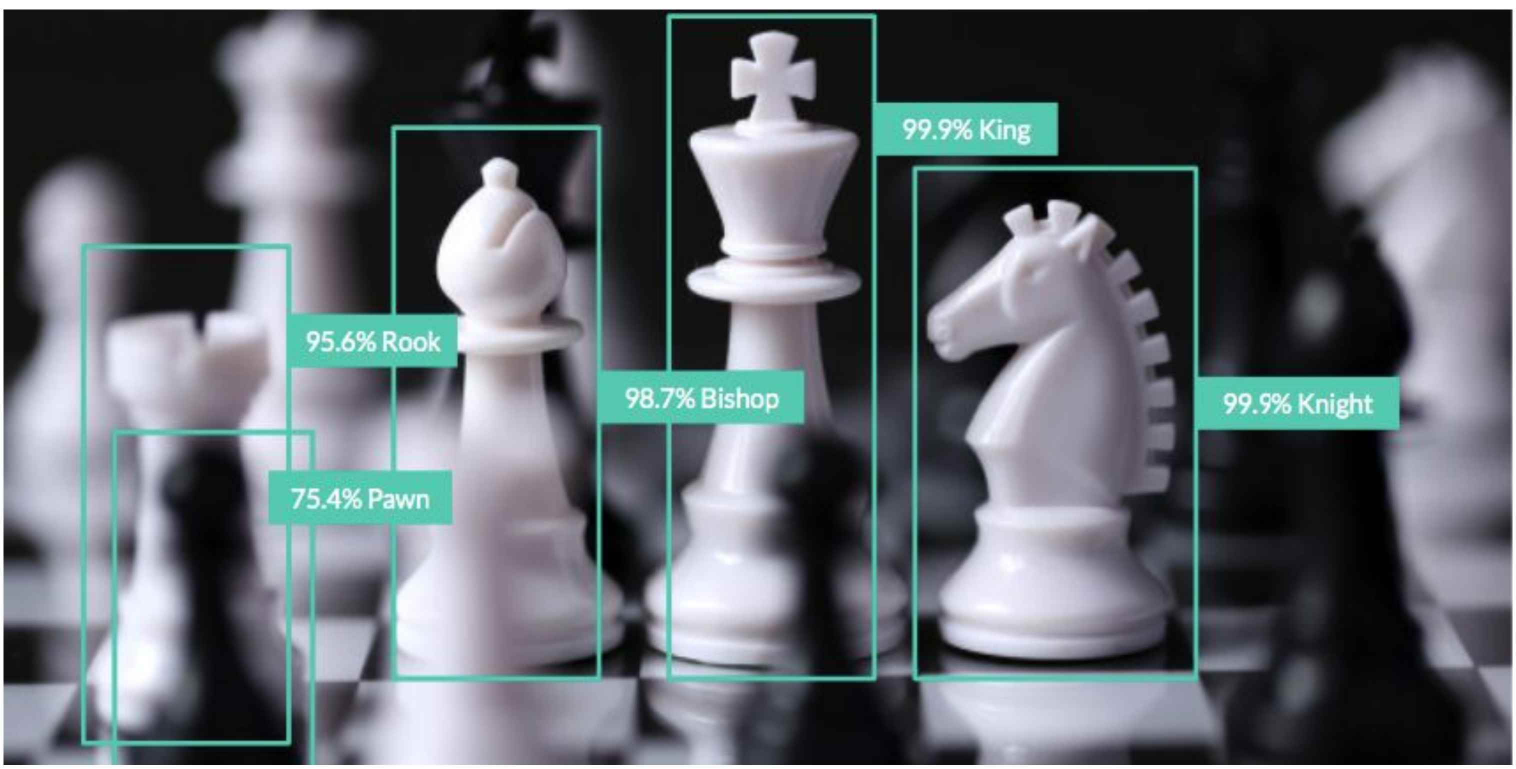
Small and medium-sized enterprise (SME) financing is a vital component of economic growth across the globe, and the need for access to capital is especially important in developing countries. However, lending markets in these countries are also the least developed, and financial institutions are often reluctant to lend money to companies without any or limited credit history. This has led to the industry naming these organizations as “thin-file” customers. The resulting credit gap that formal SMEs face is about $1.5 trillion.[1] This problem is especially acute in the Asia Pacific region in which 40% of the gap originates, totaling $600 billion.[2]

The Asian Development Bank reports that 74% of rejected trade finance transactions come from SMEs, and at least 36% of rejected trade finance may be fundable. This suggests that there is tremendous opportunity to address the underserved SME market, but how?

In this white paper we discuss how artificial intelligence (AI) and machine learning (ML) can unlock value from the treasure trove of data trapped in the databases of traders, banks, logistics companies, and others that could - in combination with alternative data sources - be algorithmically predictive in guiding risk management to unlock SME finance.

With machine learning models trained using hundreds of billions of transaction data, Flowcast delivers a game-changing approach to credit decisioning for thin-file customers, empowering financial institutions to close the significant SME finance gap.

Flowcast’s machine learning model takes a similar approach to Tensorflow’s Object Detection API. Analogous to deploying trained models capable of identifying multiple objects in a single image, Flowcast can accommodate the varying data available across companies. And, even with limited data, Flowcast’s technology can borrow from the behavioral patterns of other companies to predict outcomes, albeit with lower confidence levels.



THE SME FINANCE PROBLEM

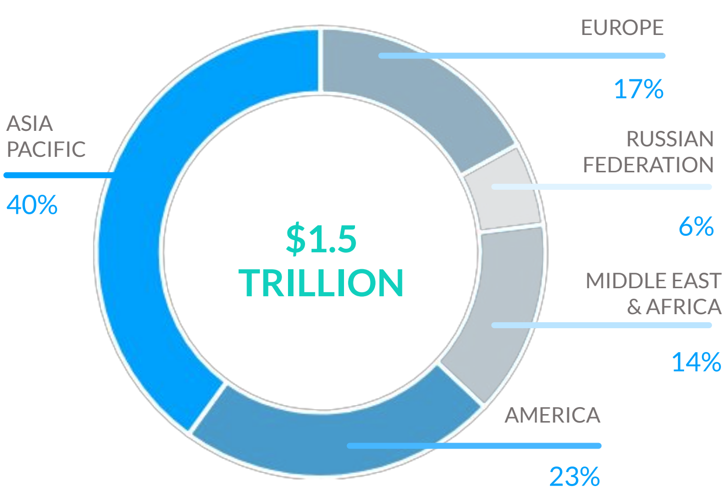
The Facts and Figures of SME Financing

According to the World Bank, "SMEs are the economic backbone of virtually every economy in the world."[3] Yet, they are struggling to find the finance needed to build their businesses.

SMEs represent more than 95 percent of registered firms worldwide, accounting for more than 50 percent of jobs, and contributing to more than 35 percent of Gross Domestic Product (GDP) in many emerging markets. SMEs generate most of the new jobs that are created, help diversify a country's economic base, and are a powerful force for innovation. Yet despite these economic, social and political benefits, SMEs remain significantly underserved by financial institutions.[1]

In investment-climate surveys around the world, the difficulty of obtaining financing is usually one of the top three constraints on doing business that are identified by SMEs – and, in several regions, access to finance is the single most important constraint. The resulting credit gap that formal SMEs face is about $1.5 trillion. When informal SMEs are taken into account, that gap widens even further, to around $2.6 trillion.[2]

GLOBAL SME FINANCE GAP



74% of rejections come from SMEs

ADB highlights that the Asia Pacific region suffers from the largest gap, accounting for 40% of the total global trade finance gap. The persistent shortfalls in the Asia Pacific region may point to the region being a hub for manufacturing supply chains. There is significant market interest from SME suppliers in emerging markets to access pre-shipment financing.

The SME Financial Challenge: A Tale of Rejection

Almost **75%** of SMEs face financial rejection from banks

Foregone trade complicates the issue

Rejection results in **2 out of 3** of SMEs unable to execute transactions

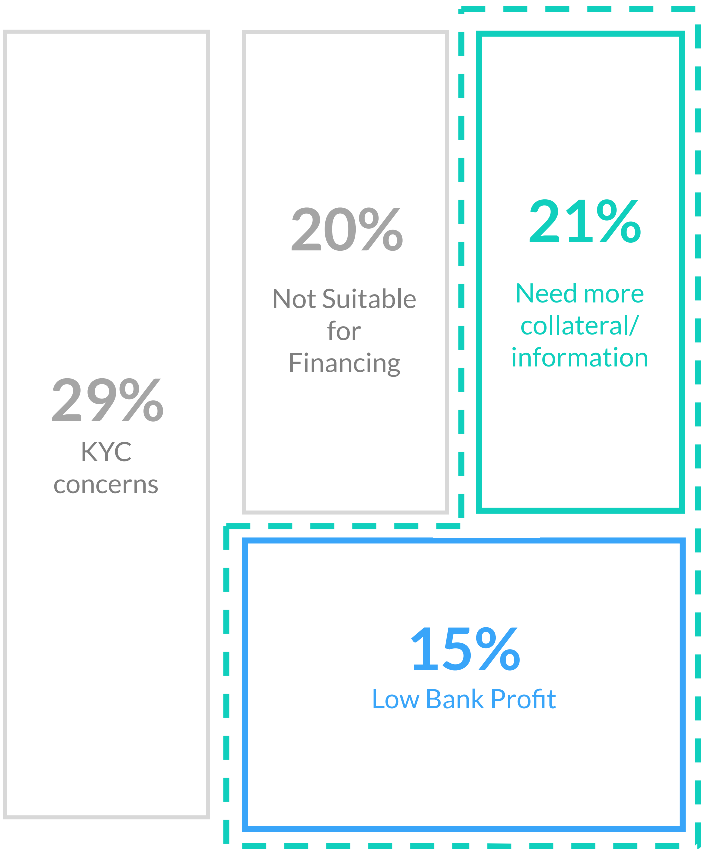
Nearly **50%** of SMEs did not look for alternative forms of finance

ADB's annual trade survey titled, "2017 Trade Finance Gaps, Growth, and Jobs Survey", reports that SMEs consistently face more difficulty accessing trade finance than large firms. The survey carried also found that banks report 74% of rejections come from SMEs.[2] A key survey outcome from these high ejection rates was foregone trade.

Why Rejection Happens

A number of variables determine a rejection outcome for SMEs applying for finance. Given 36% of rejected trade finance transactions are considered viable, these variables need to be looked at.

Four main reasons for SME Rejection:



Source: Asian Development Bank

Those SMEs falling into the first 2 types of rejections would potentially be fundable by other financial institutions such as fintech firms, which have different requirements. On the other hand, this represents a significant opportunity for lenders, who - when equipped with the right tools like AI and machine learning - can tap the underserved SME market.

*This represents **540 billion** USD of financeable opportunity*

Building a Bridge Between Bank and SME

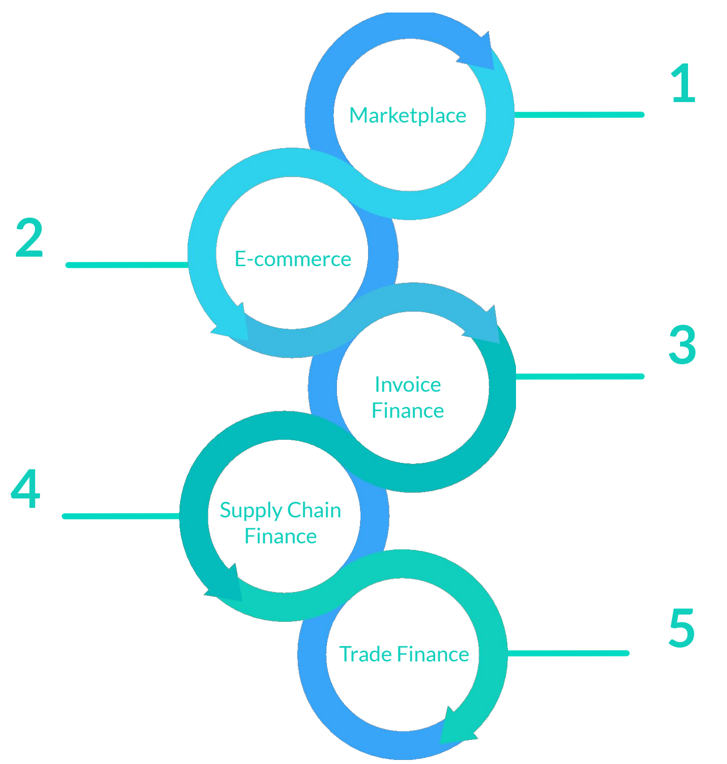
Supporting the financial needs of the SME community is crucial to not only a country's growth, but the overall global economy. Building a bridge between the bank and the SME opens up a key and loyal customer segment.

Closing the SME finance gap has a positive tri-fold effect benefiting the global economy, SMEs, and banks alike. Unfortunately, despite the benefits there remains significant systemic shortfalls inherent in the SME financing ecosystem.

INNOVATIONS IN FINANCING SMEs

A new approach is needed to ensure SMEs have the liquidity to grow their businesses. In recent years, innovative products and new business models have emerged.

Financial technology, or Fintech as it is known, has provided these core innovations. There is now the potential to significantly increase access to finance for SMEs. Five key products to provide this funding for small businesses have been identified:



Used together, employing these solutions can have enormous positive effects on an SME's balance sheet situation, leaving small businesses with more cash, improved working capital management, and more stable and secure funding.

RISK LEVEL

MARKETPLACE (PEER-TO-PEER) LENDING

Marketplace or peer-to-peer (P2P) lending platform was introduced in 2005. This is now a global market with a multitude of different business models and high projected future growth rates. This route offers lending to borrowers without going through a traditional bank.

Incumbents: Lending Club, MarketInvoice, Funding Circle

MERCHANT AND E-COMMERCE FINANCE

E-commerce platforms, payment processors and telecom companies, including Amazon, eBay or Alibaba. More and more merchants are now offering working capital and loans. Payment processors also offer similar services.

Incumbents: Amazon, Alibaba, Paypal, Square

INVOICE FINANCE

Businesses sell unpaid receivables to a third party ("factor") to improve their cash position. The receivables are bought at a discount against cash payment, along with a retainer once the customer has paid. This is called factoring and is predicted to grow at a rate of 10-12% each year. SMEs often cannot access traditional factoring that generally requires long-term, complex contracts with fixed volumes.

Incumbents: Factoring companies

SUPPLY CHAIN FINANCE

Supply chain finance (SCF) can improve an SME's working capital. In contrast to invoice finance, SCF is usually initiated by the buyer. Traditional SCF requires a high level of cooperation and integration between the smaller suppliers and buyers

Incumbents: Banks

TRADE FINANCE

It is estimated that the volume of trade finance per year is at least five times the amount of US dollars in circulation with information technology supporting the selling across borders. Online marketplaces have opened global trade to the SME. However, transactions suffer a high degree of friction. This is due to purchase and delivery often being several weeks or even months apart for custom-made products. Traditionally, banks on both sides are involved and this involves complex processes and detailed documentation.

Incumbents: Banks

Challenges in Traditional Credit Underwriting

One of a lender's most challenging task is accurately assessing the credit risk of SMEs, which presents a major obstacle to SMEs gaining access to credit.

Four issues identified are:

1. Limited and **fragmented** financial data
2. **Insufficient** risk models
3. Lengthy and **time-consuming** processes
4. Broader issues that exist internally such as the **tension** between sales and credit

Source: Moody's

Other factors having an impact include different geographies and positions in the economic and credit cycles.

The two main limitations of best practices are:

1. Many SMEs are **'credit invisible'**, that is credit scores are unavailable for them.
2. The **accuracy of scores is insufficient** in separating good borrowers from bad borrowers.

The SME Invisibility Cloak

SMEs typically are more "opaque" than large corporates due, in large part, to the lack of publicly available financial data.

Traditional credit scoring is linear, static, and one dimensional

This 'opaqueness' makes it especially difficult for lenders to determine the creditworthiness of SMEs. Often, lacking a single piece of information can prevent an application from being assessed altogether, resulting in 'credit invisible' SMEs.

Traditional credit scoring does not serve SMEs well. The method is linear, static and one-dimensional. The Altman Z-score, a technique commonly used by lenders in traditional credit scoring, is unsuitable for SMEs because it is based on a highly selective number of fields that does not take into consideration other valuable accounting and non-accounting data. If an SME lacks information in one field, it is not rated by the lender. This results in a 'credit invisible' SME. Attempts to improve the current methods of credit scoring that is accuracy and predictive for this particular segment is lacking.

As such, banks are heavily reliant on relationship-based lending when dealing with SMEs. However, relationship-based lending is not the preferred means for lenders, especially when dealing with SMEs with smaller ticket sizes and less cross-sell opportunity. Lenders need to achieve scale, and lower the cost of acquiring and underwriting credit for SMEs to make it worthwhile. New technologies such as AI and machine learning are proving to be a promising solution.

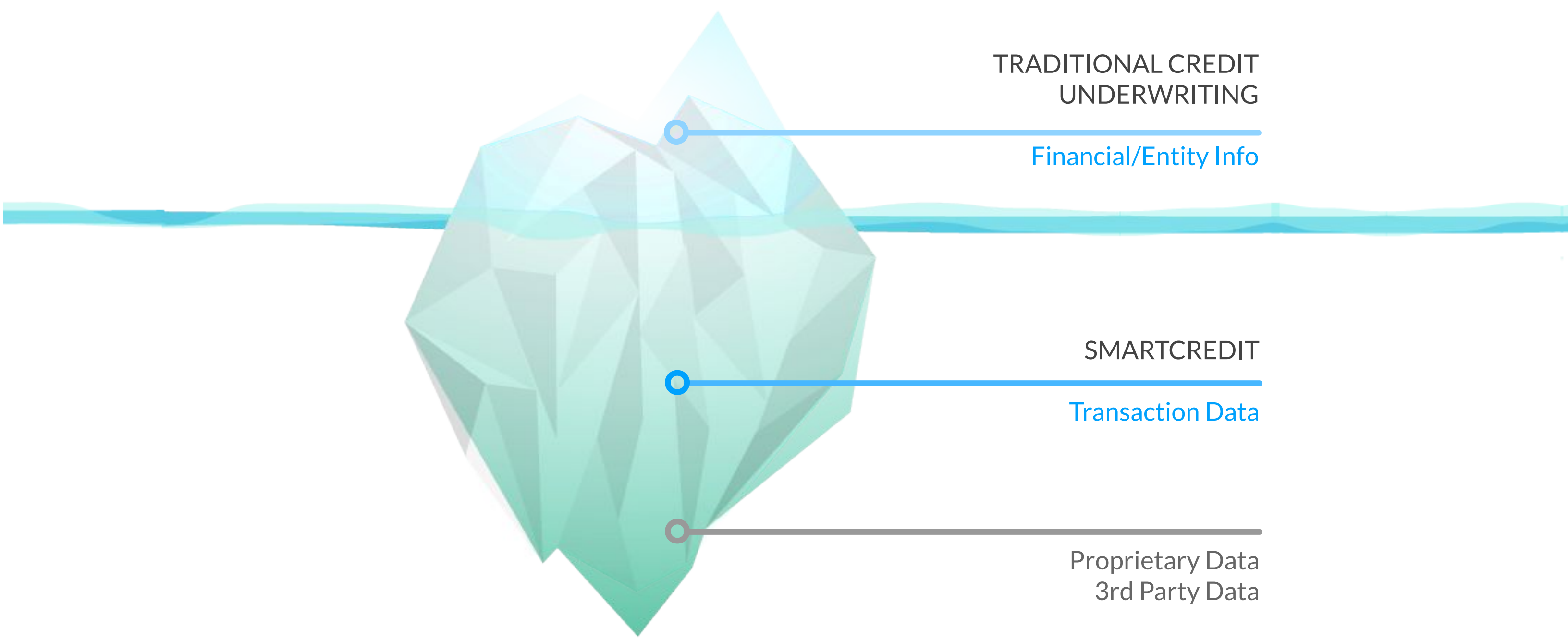
Lacking a single piece of information can prevent an application from being assessed altogether, resulting in 'credit invisible' SMEs

Alternative Data to Uncover Thin-file, Credit Invisibles

'Alternative data' is instrumental for SMEs to gain access to credit. AI and machine learning (ML) transform existing fragmented data and combine it with alternative data to produce 'Smart Data'. While we don't believe such alternative data can completely replace traditional credit data, we have seen that the combination of such distinct datasets could enhance credit assessment, powerful enough to bridge the gap for the un-banked and under-banked. Because machine learning allows for better risk separation, it has the ability to uncover many of today's credit invisible SMEs.

Differentiators and Benefits of Flowcast's Machine Learning Algorithms

As discussed throughout this white paper, access to credit is a key constraint for SMEs. The immense trade finance gap for this customer segment is caused by the difficulty of predicting business risks. Traditional scorecard methods are static and create 'blind spots'. The innovation that Flowcast has made around smart data analysis has a number of key benefits that remove these blind spots to give true visibility to thin-file, credit invisible SMEs.



Harness Untapped Data for Greater Visibility

Smartcredit's proprietary machine learning algorithms and statistical models can draw on more diverse data types, and more detailed and current company data to provide accurate and predictive credit-decisioning on thin-file, credit invisible SMEs. Smartcredit makes it possible to leverage existing fields of thin-file SMEs that are uncaptured by the more traditional scorecard methods. Flowcast does this by utilizing a wide scope set of available and emerging data sources along with our proprietary alternative data sources. This results in a 'deeper data' analysis that helps to accurately predict the creditworthiness of SMEs that were previously uncaptured by lenders.

Smartcredit leverages untapped complex data to automate credit decisioning. Our customers have greater visibility to address the underserved market.

Fact: Supply chains offer an invaluable data source

There is a tremendous amount of structured and unstructured, financial and non-financial data trapped in supply chains not being put to use. While not all data is useful, Smartcredit is able to extract value from these alternative data sources trapped in the supply chain ecosystem in order to measure performance risk, examples being dilution and late payment days.

Fact: Ensuring the use of current data improves visibility

Smartcredit uses current transaction data in addition to data those filed annually – or even less frequently – in a company's accounts, which can make those account up to two to three years out of date. Using higher frequency and granular data on companies leads to greater visibility and a more accurate picture of future risks.

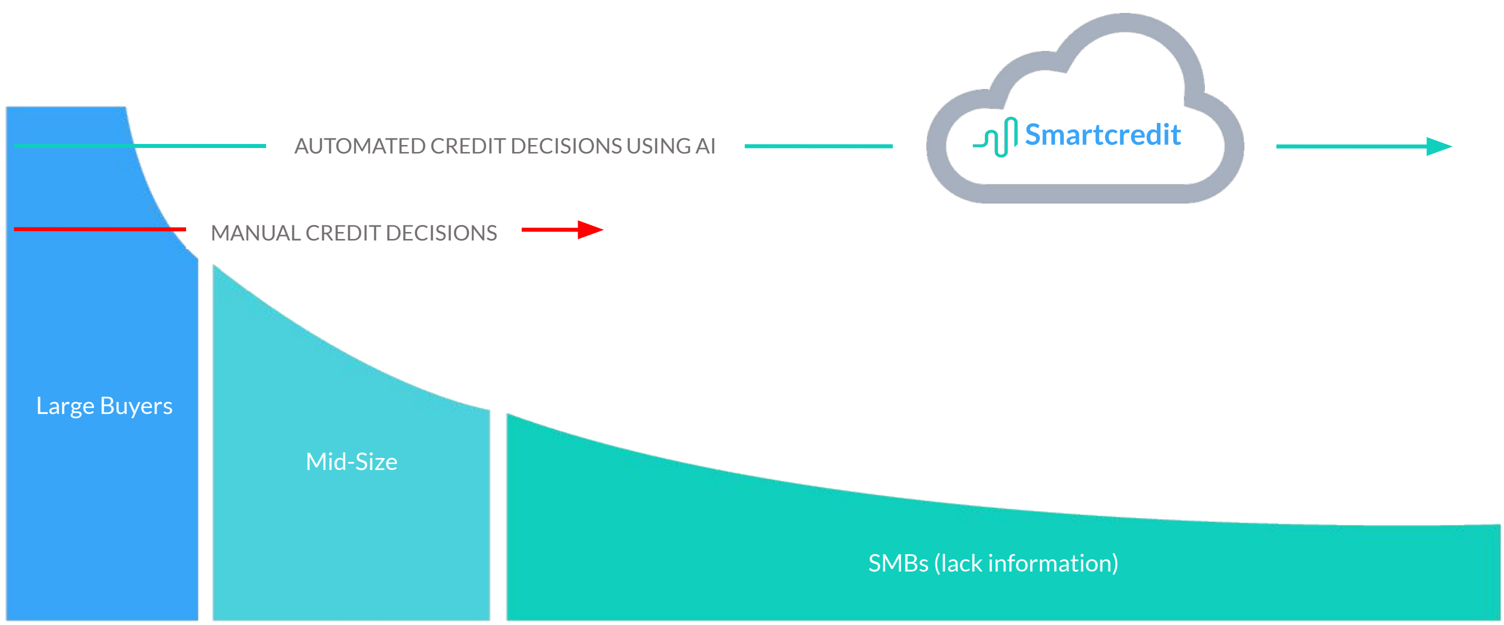
Towards Smarter Financial Inclusion

Flowcast's Smartcredit processes vast and deep data sets, uncovering thin-file, credit invisible SMEs. This method of credit scoring captures a far greater number of SMEs than traditional scorecard methods, because it allows for thousands of data points to be analyzed and at a far faster rate, mere minutes, than a human underwriter is capable of manually. Smartcredit helps lenders achieve the scale that is needed to serve the SME segment, resulting in more financial inclusion.

Measure Performance Risk

Flowcast leverages its patented machine learning algorithms to create high-performing predictive models that reduce risk and unlock credit to businesses. Our models have demonstrated high accuracy in predicting the ability of a business to repay its loans, the likelihood of dilution, and the risk of delinquency. Our API-based ML model, actively predicts the changes in behaviors based on the activities along the supply chain ecosystem.

Supply chains contain a treasure trove of data, including payment performance, order pipeline, product information, and end user data that can reveal creditworthy insights of those involved: both buyer and supplier. Smartcredit processes transaction data trapped in the supply chain ecosystem and combines it with proprietary alternative data to measure performance risk.



AI and machine learning models make it possible to efficiently surface hidden patterns between various risk levels of all parties involved in the supply chain ecosystem. A machine learning approach is an extremely powerful tool for lenders, giving them visibility beyond default risk into performance risk.

Smartcredit empowers lenders with the ability to predict:

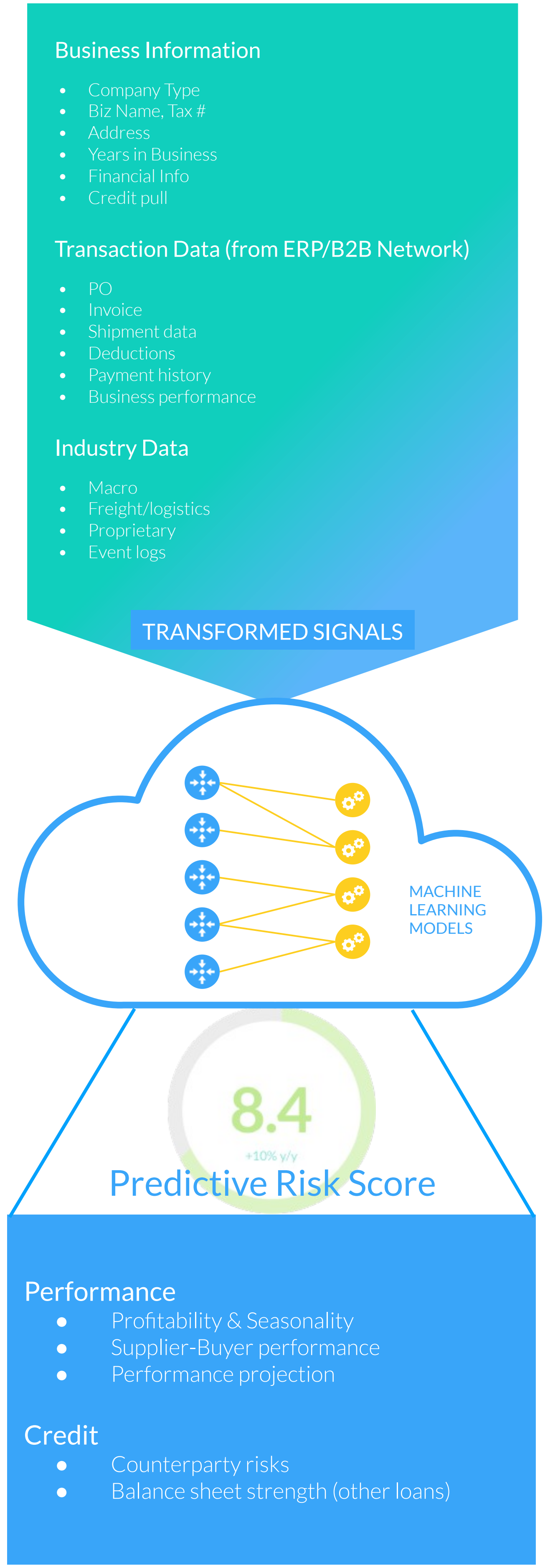
1. **Dilution** of incoming and open transactions
2. **Lateness** of payments

These predictions, are in turn, calculated into a single numeric score for each client to make the prediction 'actionable'. The solution is based on a machine learning algorithm trained with millions of global transactions using data across multiple products. The algorithm and predictions Smartcredit generates are agnostic to any specific workflow and are expected to be implementable in connection with existing or future credit or operational workflows.

Self-Improving Algorithms

AI and machine learning enables the generation of further efficiencies, as it develops patterns from its own history. The self-improving algorithm uncovers intricacies of SME behavior, allowing for a far more accurate and finely-tuned means of assessing the true creditworthiness of SMEs.

A machine learning approach is an extremely powerful tool for lenders, giving them visibility beyond default risk into performance risk



Business Benefits of Using Machine Learning

Smartcredit's proprietary machine learning algorithms enable predictive modeling in credit scoring. AI and machine learning evaluates data at a larger scale and aggregates the data through wider and more up-to-date channels. Smartcredit also leverages alternative data sources to measure performance risk. The combination of measuring credit risk with performance risk is much more robust that the traditional scorecard method. Smartcredit delivers a far more accurate and fine-tuned representation of the creditworthiness of borrowers, resulting in fewer loan rejection rates and a healthier portfolio for lenders.

REDUCE THE COST AND TIME OF MAKING A LOAN

MAKE MORE LOANS

CONTROL RISK MORE EFFECTIVELY

REMOVE HUMAN BIAS FROM THE LENDING DECISION

FOCUS ON ASSESSING QUESTIONABLE LOANS

"AI tools provide the most compelling and straight forward way to increase returns than virtually any other technology available in the working capital space" - Tom McCabe, US Country Head of DBS Bank

The application of Smartcredit's machine learning to financial risk offers three major benefits:

1. **Speed** - Data-driven intelligence provides quick and confident credit decisions
2. Address the **long-tail** - Predicts creditworthiness for the thin-file, credit invisible segment
3. **Accuracy** - Models have been tested over extensive periods of time and trained over vast amount of data

The use of Smartcredit enables banks to:

1. Support Growth:

- Identifies new financing opportunities
- Supporting decisions to extend existing credit

2. Optimize Products:

- Identifies low-risk opportunities for tighter limits or enhanced credit line utilization

3. Protect Against Risk and Comply with Regulations

- Objective, unbiased predictions
- Provides an additional model for monitoring risk

Smartcredit Technology Highlights

This overview gives you a snapshot for the underlying technology behind Flowcast's Smartcredit platform.

Our model is:

1 **Explainable:**

Unveiling the 'Black Box' is key for AI adoption in credit risk

2 **Generalizable:**

Robust model that can scale across a broad and diverse set of entities

3 **Proven:**

Trained with over hundreds of millions of transaction data - available for internal validation

Ease of Use and Extensibility

Flowcast has developed an enterprise scale machine learning platform that provides a unique and effective approach to credit decisions. Our API-based solution allows financial institutions to deploy machine learning models that can turn complex data and insights into actions without writing a single line of code.

Always Improving

Smartcredit allows risk professionals to gain the benefits of transfer learnings. It was designed to help rapidly analyze complex datasets to drive more accurate and auditable risk models. Our machine learning solution allows access to insights that are captured through many iterations and experience working with risk professionals around the world, uncovering intricacies of supplier and buyer behavioral dynamics.

Explainable

There are regulatory requirements to be able to explain, in human terms, why certain decisions were reached. The ability to explain the predictive output is equally important to the accuracy of the prediction itself. The rise of complex machine learning models makes explainability a key element in AI.

Validated and Seamless

We provide analytics tools to make validation and monitoring seamless. We require our models to perform well in a real-world setting. This means that we carefully assess the out-of-sample and out-of-time sampling in our validation set and compare that against the in-time training set. It is important to have accurate assessments of future performance in production.

API First

Customers can invoke an API call in real-time to get predictive results. These can then be integrated into existing ERP rules engines and workflow tools. The machine learning platform can be deployed in public or private cloud infrastructures.

CASE STUDY

CLIENT PROFILE:

A LEADING GLOBAL BANK

RESULT:

Successful transformation of trade finance credit risk using Flowcast's machine learning solution and methodology. The Bank uses Flowcast's ML platform for trade finance to better serve their existing and new clients. The Flowcast solution offers the Bank a means to better understand their customer needs and gives greater visibility of the ecosystem risks. The use of the Flowcast ML platform eliminates the rigid limitations placed by conventional risk assessment.

Unlocking the Value of Transaction Data

The effectiveness of a conventional credit model hinges heavily on the quality of the data. Indeed, conventional underwriting primarily focuses on 'thick-file' customers. Such customers have sufficient credit information, such as client identity, credit account information from reliable 3rd party providers, and financial statements that can be used in a scorecard method to measure probability of defaults (PD). In contrast, a 'thin-file' customer (an individual or business) has zero or limited credit information. This cohort of clients will not be considered in conventional credit assessment.

However, Flowcast can combine the limited credit information of these 'thin-file' customers with unconventional data such as transaction data to achieve results. In our data discovery journey, we processed a large variety of data sources across different client bases; this included, global financial institutions, credit insurance providers, factoring companies, B2B networks, payment service providers, and corporations. In our example of working with the global bank, we successfully harnessed a treasure trove of internal transaction data that were previously underutilized for credit risk assessment. These transaction data (PO, goods receipt, invoice, payments) was traditionally not used in the conventional credit model.

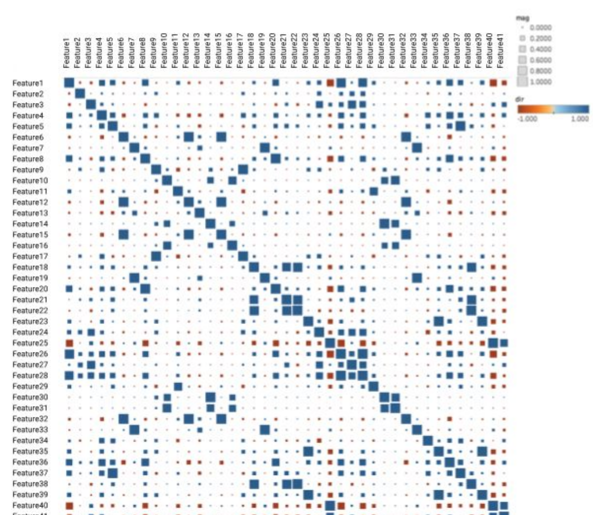
Such historical transaction data served as a guide to suppliers and buyers' performances. However, to derive any meaningful insights, we had to apply advanced machine learning methodology in order to capture the non-linear relationships across the dataset.

In addition to transaction data, we incorporated other data sources including public data (government records, business licenses, freight data, industry indices, country related data, etc.) as well as other proprietary data sources to enhance the machine learning model's predictive power.

Feature Generation

One of the key elements of the Flowcast ML platform is in the feature generation module. Our insights are captured through many iterations based on experience working with various stakeholders in credit risks, in-country relationship managers, trade finance product managers across different financial institutions, and corporations. Such insights around the intricacies of supplier and buyer behavioral dynamics are captured in our feature generation.

For example, certain product categories during a seasonal time in a specific region have a higher tendency for dilutions. We capture and scale up these insights into time-series features that would be used in model development. We then perform a set of analyses to gain further insights into the relationships of a set of feature attributes and performance. In some situations, we apply 'binning' to convert numerical values into categorical variables when we are dealing with skewed or rare occurrences.



Flowcast Machine Learning Algorithm Selection

There are a number of algorithm choices supported by Flowcast. Our ML solution stack includes different ML algorithms, including Penalized Regression, Random Forest, Boosted Trees, and so on.

Random Forest (RF) for example, has certain key benefits including ease of use; lack of significant hyperparameters; lack of fitting bias; ability to allow large amounts of features as inputs; and, resilience to outliers and feature collinearity.

The following table outline the pros and cons of three common ML algorithms:

Algorithm	Benefits	Drawbacks
Penalized Regression	Easy to interpret & understand Penalization allows for more features than standard regression	Must hand-code variable interactions & nonlinearities Saturates with too many variables
Random Forest	Almost no tuning parameters Variable interactions & nonlinearities fundamental Little variable prep necessary	Trees are independent so count on averaging for predictive power Cannot easily be explained
Boosted Trees	Iterative algorithm that identifies and targets errors Automatically detects interactions & nonlinearities State-of-the-art performance across many data science problems	Difficult & time-consuming to train and tune Can easily latch on to biases or errors in data Opaque to understand

In our bank case study, we explored the hyperparameter space to decrease overfitting of data, including setting the minimum number of samples per leaf to a value > 1. We choose several hundred trees to balance generalizability with performance. Since our goal is automated retraining, we want to allow ample complexity to handle future unforeseen interactions.

Generalizability of Model

The ML solution we developed in trade finance reinforces our single model approach without the need to develop individual models for each seller. The current solution has a monolithic model applied to thousands of suppliers, comprised of over 10 million transactional data in aggregate counts. The supplier base is also very broad and diverse, spanning across many industries, countries, sizes, and length of history. The algorithm leverages the structured data across those suppliers to produce a single unbiased model with predictive results applicable to existing and new clients, regardless of their size (from SME to large corporates).

Reducing Bias for High Prediction Performance

A major challenge in a ML model is controlling the tradeoff between bias and variance (we want low bias low variance). For example, having distinct individual models for each client could result in high bias and low variance. From a machine-learning perspective, it is beneficial to leverage data across all suppliers to build one model. The algorithm is unbiased and develops unseen connections across different types of clients to produce a high predictive performance.

A single model is much more robust and generalizable if we can include more buyers and sellers. The generated features become more important and the model will depend less on ID fields particular to a single dataset. ML models benefit from diversity in the inputs, and pooling across many sellers and buyers will allow the algorithms to lower variance, rather than concentrating on the particular details of a single client that may not apply to others, which leads to higher bias.

Validation of Machine Learning Model

Under banking regulatory guidance, all models must be validated with a set of processes to verify the model performance under business objectives and design. The model must be auditable and verifiable. The ability to validate the results is a significant part of our ML platform. It is important to have accurate assessments of future performance in production.

The validation process has two components:

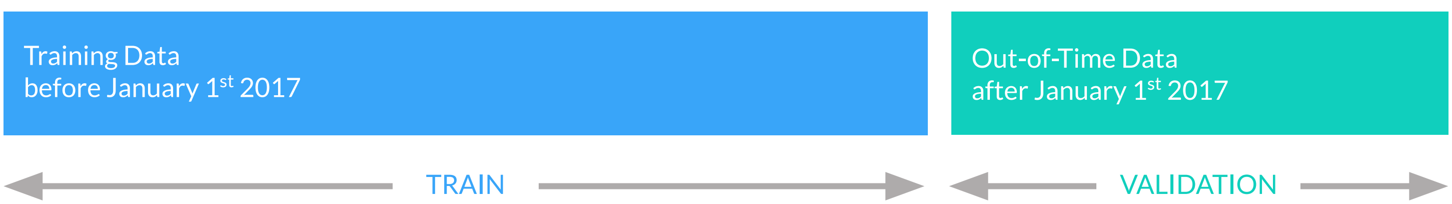
1. Train set: The Random Forest algorithm data prior to the 'cutoff date' is used for the model building process. For each tree built, the data is randomly sampled using bootstrap independently of all other trees. These trees form the ensemble Random Forest ML model.

2. Out-of-Time (OOT) Validation set: The most recent data (6-9 months) is set aside from the Train set. The performance on this data is not investigated until the model is complete. Because it includes new time periods that are 'unseen' by the model, this is the best data on which to determine the model's expected production accuracy. This is the data on which we generally report accuracies and define thresholds/cutoffs for the prediction buckets.

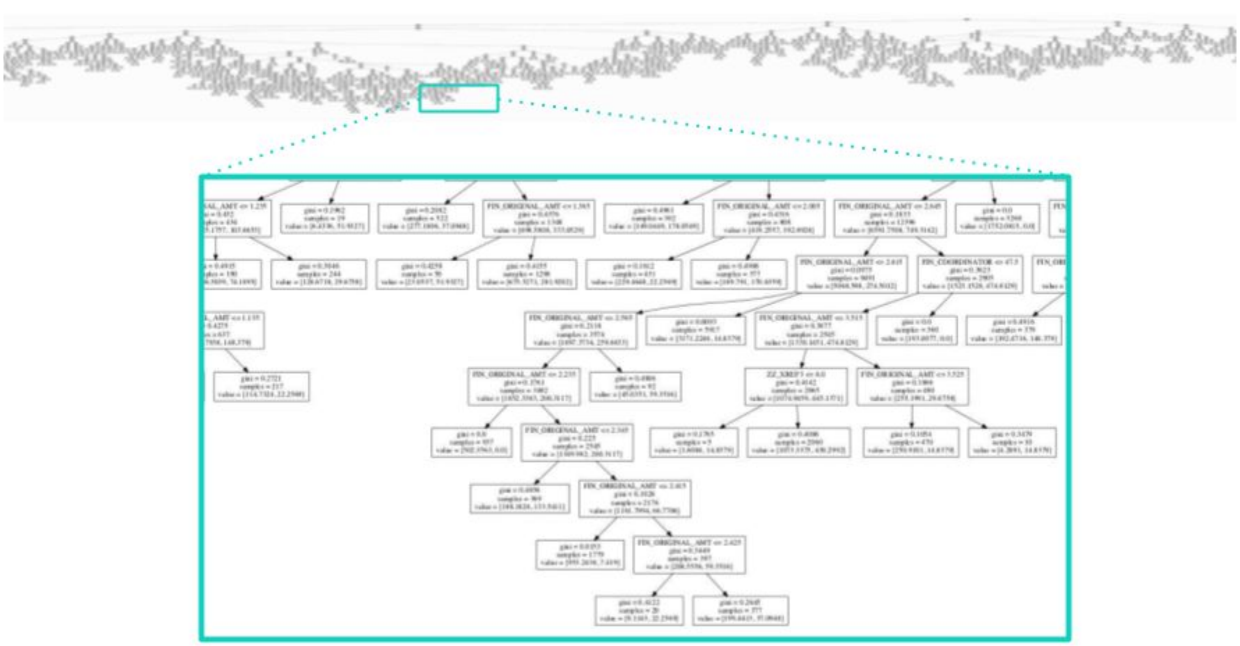
We require our models to perform well in a real-world setting. This means that we carefully assess the out-of-time sampling in our validation set and compare that against the train set. We also assess the out-of-sample (eg., setting aside a sample set of supplier history completely) as a further comparison. As an example, the model performance in trade finance confirms how little we overfit (89% accuracy between in-time test vs. 85% accuracy in out-of-time validation). In addition, our client heavily scrutinizes our model to ensure its performance on a go-forward basis. This requires a shadowing process to let the models run for a period of time and measure the on-going performance.

The Replacement Question

One misconception in machine learning is that it could replace risk officers with decades of experience and specific country and domain knowledge. Machine learning doesn't give us automated risk experts, just like Excel doesn't give us artificial accountants. Rather, it is trained to replicate the insights across infinite cases that humans cannot possibly scale to.



EXAMPLE: Cutoff date is set at January 1st , 2017



Prediction Accuracy

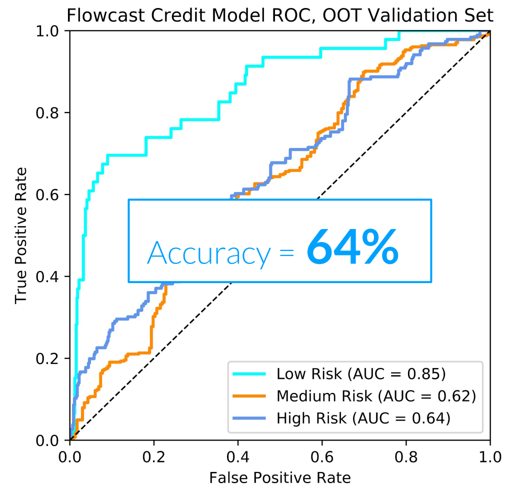
While the model predicts a continuous variable, the bucketed classification performance is illustrative. We set the outcome ranges of 0 - 0.33 for low risk, 0.33 - 0.66 for medium risk, and 0.66 - 1 for high risk. The OOT confusion matrix is a matrix of predicted versus actual labels and is used to determine real-world performance. False negatives appear in the lower left quadrants while false positives appear in the upper right. The accuracy is the sum of the diagonal elements. This is 85%, as shown below using the random forest algorithm. We want to minimize the false negative error rate, which are medium/high risk clients that we would predict as low risk. In the OOT sample this error rate is 0.5%.

PREDICTION

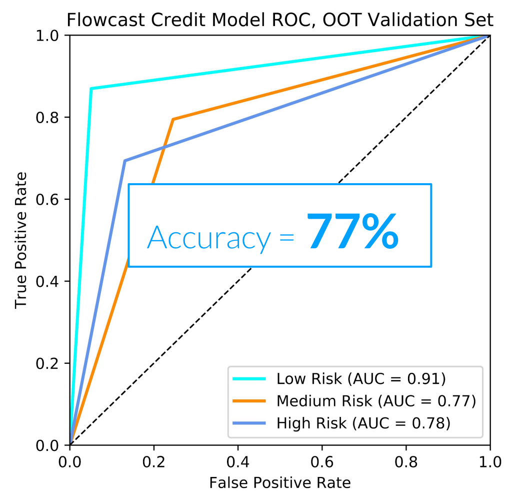
ACTUAL	Low Risk	Medium Risk	High Risk
Low Risk	5.5%	1.7%	0.0%
Medium Risk	0.5%	60.4%	2.7%
High Risk	0.0%	10.1%	19.2%

The receiver operating characteristic (ROC) curve is a benchmark of model performance. It is useful when the outcome labels are unbalanced, which they are in these cases. The plot is the false positive rate (sum of false positive cases / sum of actual negative cases) versus true positive rate (sum of true positive cases / sum of actual positive cases). Each point in the ROC curve represents one threshold in the confusion matrix. The diagonal line represents the baseline performance of the average prevalence of outcome labels. Perfect models lay on the y = 1.0 line. In the multi-class case, each outcome is tested separately. The area under the curve (AUC) is representative of how performant the model is per outcome. As shown, the random forest significantly outperforms the other algorithms such as logistic regression and the decision tree.

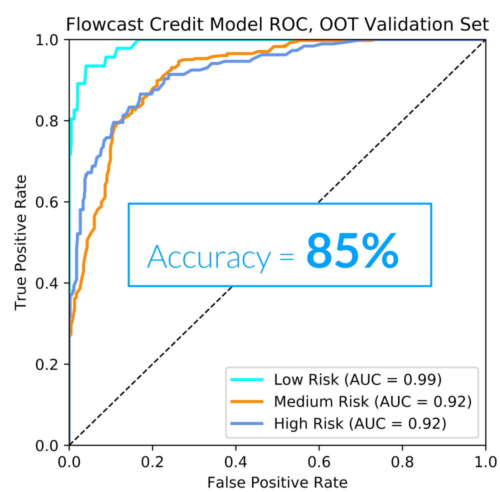
LOGISTIC REGRESSION



DECISION TREE



RANDOM FOREST



Explainability: Unveiling the "Black Box"

The rise of complex ML models makes explainability a key topic in AI. In lending, there are regulatory requirements to be able to explain, in human terms, why certain decisions were reached. The ability to explain the predictive output is equally important to the accuracy of the prediction itself. However, given the complexity with non-parametric machine learning model, it is often difficult to uncover features that describe the prediction.

Consider a trade transaction that could have hundreds of attributes, from shipment integrity to diversity of the customer base. Visualizing the complex relationships and patterns within the supply chain can be quite difficult. One technique that we use to make the machine learning model transparent is 'local-interpretable-model-agnostic explanations' (LIME).

LIME segments the dataset in smaller pieces and time intervals. It then generates an interpretable model for each segmented data in which becomes a proxy. Flowcast is making advances in furthering LIME deployment with the goal of AI adoption in credit decision making.

FINAL WORDS

The market for SME finance is untapped because we are unable to decloak the data needed to inform us of credible risks. Using conventional methods to determine finance decisions are proving too restrictive and can give us an obfuscated view of an SME's ability to repay finance. To remove the invisibility cloak around SME data, we need to extend our reach out to new forms of data, including the supply chain. This creates a more transparent, real-time, and holistic view of a company's risk-profile.

In this paper, we present the results of our proprietary credit risk models produced from a data-driven approach through machine learning. We present our methodology to produce machine learning models that are highly performant and equally important - explainable. Our model performs with 85% accuracy with false negative error rate at < 0.5%. We strive to produce highly performant models. However, the tradeoff with high performance is model obscurity. Therefore, we provide interpretable explanations along with each prediction to provide insight as to how the model behaves in different data regimes.

Smartcredit is a powerful tool for lenders who wish to leverage AI. It opens up your organization to the accurate use of non-traditional data sources to gain insight. The platform provides all of the requisites needed to unlock transaction data, giving you the tools to build SME financing models based on intelligent decisions using the power of machine learning.

We hope you enjoyed learning more about how artificial intelligence and machine learning are transforming credit decisioning for thin-file customers, empowering financial institutions to address a largely underserved market and close the significant SME finance gap.

If you wish to know more about us, feel free to contact us at info@flowcast.ai

REFERENCES

1. World Bank, What's Happening in the Missing Middle?: https://tinyurl.com/ycr62pkn
2. Asian Development Bank, 2017 Trade Finance Gaps, Growth, and Jobs Survey: https://tinyurl.com/y7zs95qm
3. World Bank, Addressing the SME Finance Problem: https://tinyurl.com/y7zs95qm
4. Moody's: https://www.moodys.com

FLOWCAST

About Us

We started Flowcast with the mission to help the underserved market gain access to capital. We had the thesis that AI/machine learning can harness the trove of data trapped in databases of traders, banks, logistics companies and others that could - in combination with alternative data sources - be algorithmically predictive in guiding risk management to unlock SME finance. We want to empower Global 500 corporations, financial institutions, and insurance providers around the world to transform the way credit assessment is done. Our core team of Postdoc PhD data scientists brings deep domain expertise in credit risk, financial services and AI/machine learning.