THE POWER OF STATISTICAL LEARNING
CUTTING-EDGE ANALYTICS FOR COMMERCIAL CREDIT RISK MANAGEMENT

AUTHORS:
Ugur Koyluoglu
Attilio Meucci
Gokce Ozcan
Simon Schwendner
Kirill Skok
Dan Wang
THE POWER OF STATISTICAL LEARNING

CUTTING-EDGE Analytics FOR COMMERCIAL CREDIT RISK MANAGEMENT
# TABLE OF CONTENTS

**EXECUTIVE SUMMARY**  

**ADVANCED ANALYTICS FOR COMMERCIAL CREDIT RISK MANAGEMENT**  

UNDERSTAND AND MANAGE RISK BETTER  
MAKE BETTER USE OF EXISTING DATA SOURCES  
MAKE USE OF VARIED NEW DATA  
INCORPORATE DATA THAT IS UPDATED IN REAL-TIME  

**CHALLENGES AND POTENTIAL PITFALLS**  

GOOD DATA – A MUST HAVE  
INSIGHTS DEPENDENT ON AVAILABLE DATA SIGNALS  
COMPLEXITY AND OPAQUENESS  
RISK OF OVERFITTING  
NON-TRIVIAL VALIDATION, GOVERNANCE AND COMPUTING  

**WHERE TO START: LEARNING FROM STATISTICAL LEARNING**  

LEADING-EDGE USERS – READY TO MAKE THE MOVE  
IF YOUR INSTITUTION HASN’T STARTED YET – HERE’S HOW TO ADAPT  
ENHANCE EXISTING MODELS AND CAPABILITIES  
DEVELOP USE CASES A STEP REMOVED FROM DECISION MAKING  
BRING SKEPTICAL STAKEHOLDERS ON BOARD  

**CONCLUSION**
EXECUTIVE SUMMARY

It’s a new frontier for commercial credit risk management – as the advancement in applied statistical learning techniques and data emerge as a game changer.

Enhanced risk models can improve credit underwriting and help monitor performance, sometimes dramatically. Institutions can build increasingly sophisticated models and algorithms, allowing them to learn from data better and faster, manage credit risk more precisely, and proactively make stronger business decisions.

While these techniques have significant benefits, they come with certain challenges and potential pitfalls, testing cultural readiness. Education and enrollment of senior stakeholders are essential for future investment and implementation. Complexity and opaqueness of these models make it harder for stakeholders, including modelers, validators and users, to grasp the embedded intuition.

To remain competitive, financial institutions often need an arsenal of advanced analytics techniques, strategies and innovative datasets – and a solid approach to lay the groundwork for the path forward.

ADVANCED CREDIT RISK MODELS – A NEW BREED TO MAKE STRONGER BUSINESS DECISIONS

For decades, banks have developed and used credit risk frameworks, including probability of default (PD), loss given default (LGD), and exposure at default (EAD) models, which relied on financial ratios, market information and structured qualitative assessments. Today, we are on the cusp of a new breed of credit risk models that incorporate signals from unstructured data, and create better and scalable risk measures to support commercial credit decisions.

This is made possible by continued advancements in computing power, statistical learning methodologies and natural language processing, all of which have accelerated large-scale data analysis, and enabled the availability of rich alternative content sets based on unstructured data.

WHAT YOU’LL LEARN MORE ABOUT

Our paper focuses on applications of statistical learning in commercial credit, covering wholesale lending, corporate, middle market and small and medium enterprise (SME) segments. We take a deep dive into the power of statistical learning, presenting our observations, research and insights through advising clients, and discussing the benefits that advanced credit analytics offer to financial institutions. Whether you’re a chief credit officer looking to develop or improve existing capabilities, or a financial modeler or data scientist

---

1 Statistical learning is a framework in which data is used to determine the functional form of a predictive relationship. The functional forms range from familiar regression choices to decision trees, tree ensembles, neural networks, and other techniques.
looking to implement an advanced analytics strategy, our paper presents the key information to get you started.

Currently, leading banks in this space are developing proof-of-concept and/or pilot versions of advanced credit risk models, and are facing many cultural, regulatory, governance and infrastructure challenges. These models are being used to replace, enhance, or act as a challenger to existing models for default or downgrade predictions.

**Advanced models are superior to traditional approaches in multiple ways and can:**

- **Make better use of existing data sources** such as an institution’s financials, for example, by effectively uncovering nonlinear relationships.
- **Make use of varied new data**, such as unstructured text data found in news and social media channels.
- **Incorporate data that is updated in real time**, which allows analysis to be timely and avoids stale signals.

**OUR FINDINGS AND INSIGHTS FOR THE PATH FORWARD**

While the innovation continues at full speed, new governance discussions have emerged in banks between model developers and the independent model risk management and audit functions: *For these advanced approaches, how should the model risk be governed where the degrees of freedom sometimes go beyond the human capacity to fully comprehend?*

Our findings and insights are based on a combination of Oliver Wyman experience advising clients, and our observations and research in this field.

- **We present what banks can learn from experimentation with statistical learning**, and provide real examples to help improve underwriting and portfolio monitoring decisions.
- **We walk through an exploration of advanced algorithm approaches**, explain the causalities, and show how to more effectively extract deeper signals.
- **We offer lessons learned**, including the challenges and potential pitfalls we have experienced in helping institutions develop the first generation of advanced models.

If your institution is ready, we recommend that you deeply engage your teams with the workings and features of these new models and study the characteristics of calibration data, including the embedded signals and biases. If your institution is hesitating, you might want to consider a pragmatic “middle path” approach which guides institutions toward implementing intelligent changes to traditional techniques without abandoning them. This approach directs developers to understand the structures, transformations, algorithms and the data profoundly to lay the groundwork for a fuller conversion to modern techniques.

We hope that the insights shared help the commercial credit community to advance faster in statistical learning, and offers commercial credit leaders and analysts a practical path to unlock the value for better credit risk management.
ADVANCED ANALYTICS FOR COMMERCIAL CREDIT RISK MANAGEMENT

Emerging technologies, sophisticated models and algorithms now offer institutions faster and better ways to learn from data, manage credit risk, and proactively make stronger business decisions.

UNDERSTAND AND MANAGE RISK BETTER

Given the advances in the underlying methodologies and computing power, we now have the means to unlock relevant commercial credit information—from alternative content datasets such as company announcements and commercial transaction records to news, social media conversations, and satellite imagery.

Leading banks have already developed proof-of-concept and/or pilot versions for a new breed of credit risk models—with stronger explanatory power and more dynamically updated information regarding creditworthiness. Models are based on statistical learning techniques and apply a combination of structured and unstructured data, promising a profound impact, particularly for portfolio monitoring. This new wave of development arrived in credit risk analytics decades¹ into banking industry-wide development and use of PD, LGD, EAD and credit portfolio models based on financial ratios, market information and structured qualitative assessments.

BACKGROUND

Quantitative hedge funds, asset managers and fintech companies embraced such capabilities and structured their business models around them many years ago.

In banking, advanced analytical techniques made headway mostly in fraud monitoring, compliance and employee surveillance as well as customer service automation.² However, until now, we have seen these capabilities lag in application to commercial credit.


² Please see Oliver Wyman publication, “Why Wall Street needs to make investing in machine learning a higher priority,” for a survey of approaches used in banking.
Through our client work and research, we have observed advances in credit risk data and methodologies.

**This activity initially focused on retail lending rather than on commercial lending—due to three driving reasons.**

1. Retail credit decisions are typically much more automated, with less individual business judgment overlaid on model outputs.
2. In retail credit, much larger datasets have made it easier to test novel hypotheses, contributing to the development of the sophisticated technical infrastructure that enables analysis of the data.
3. Due to limited historical observations of defaults and downgrades, commercial credit has always been a blend of art and science.

**BENEFITS OF STATISTICAL LEARNING IN COMMERCIAL CREDIT RISK MANAGEMENT**

Recently, despite the challenges involved, leading commercial credit practitioners have started to unlock value from unstructured data via the use of statistical learning.

The benefits of statistical learning are applicable for:

**Underwriting.** These new models offer better differentiation between marginal credit cases. The models capture a broader range of information including nonlinearities and additional factors, and identify stronger statistical relationships in unstructured data in a systematic fashion that, in the past, would have been only considered qualitatively at a high level.

**Portfolio monitoring.** Periodic releases of backward-looking financials are supplemented by additional data, including real-time indicators of credit quality changes. For example, news and announcements both drive and reflect changes in the sentiment toward an obligor’s credit worthiness. Statistical learning captures this information and reaps the benefits from more data, dynamic updates and timely execution—both directly (by including dynamic information in credit models alongside financials) and indirectly (by using the information to generate warning signals for credit analysts).

While the applied approaches come with a host of analytics and large-scale data manipulations which take a lot of effort to review and challenge, we believe the “features may be discovered in a black box, but the strategy is developed in a white box.”
HOW TO APPLY STATISTICAL LEARNING IN COMMERCIAL CREDIT RISK

The quantitative techniques discussed in this paper under the umbrella of “statistical learning” have one thing in common—they are, ultimately, used to fit an equation to describe the relationship between an outcome of interest (for example, the default or downgrade of a commercial obligor) and a set of input data hypothesized to be predictive of the outcome (for example, the obligor’s most recent financials, a large legal settlement announcement, or a sudden news development such as a management scandal).

Statistical learning techniques, typically referred to as “supervised learning,” are used to train a credit model based on data with labeled outcomes of interest, such as the “goods and bads,” representing no-default and default. Unsupervised learning, which relies on finding patterns in unlabeled data, is another area of great promise, but is outside of the scope of this paper.

With increased computing power, statistical learning techniques offer three benefits to data scientists and credit risk modelers.

THE ABILITY TO

1. Make better use of existing data sources such as an institution’s financials, for example, by effectively uncovering nonlinear relationships.
2. Make use of varied new data, such as unstructured text data found in news and social media channels.
3. Incorporate data that is updated in real-time, which allows analysis to be timely and avoids stale signals.

The rest of this section explains how these benefits can be realized using statistical learning techniques.

MAKE BETTER USE OF EXISTING DATA SOURCES

In practice, most traditional probability of default (PD) models are based on regressions which assume a simple relationship between a small set of input factors and the output of interest. These traditional regression models are constrained by the number and type of variables they can effectively capture. Statistical learning algorithms have the flexibility to capture a wide range of variables, non-linear relationships and interaction terms.

 Much of our discussion of the mathematical elements of statistical learning owes to the insights made by Trevor Hastie, Robert Tibshirani, and Jerome Friedman in The Elements of Statistical Learning (2009).
Statistical learning algorithms can:

<table>
<thead>
<tr>
<th>CAPTURE NON-LINEAR RELATIONSHIPS</th>
<th>INCORPORATE CROSS-VARIABLE INTERACTIONS</th>
<th>INCLUDE A LARGER NUMBER OF VARIABLES</th>
<th>PERFORM BETTER EVEN WITH “DIRTY” DATASETS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provide a flexible approach with potential segments to the inputs.</td>
<td>Easily capture a variable’s impact in relation to other variables.</td>
<td>Use a larger number of variables simultaneously than traditional regressions.</td>
<td>Show greater ability to capture variable interactions even with local data pollution (missing values).</td>
</tr>
</tbody>
</table>

**CAPTURE NON-LINEAR RELATIONSHIPS**

The commercial modeling dataset in Exhibit 1 reveals a strong linear relationship between default risk and the profit margin for companies with negative profitability. However, once a company becomes profitable, we observe a jump to a much lower risk level—but beyond this there is no clear relationship.

A flexible three segment approach has better explanatory power as compared to this one variable approach, which is defined the same way across all levels of profit margin.

**INCORPORATE CROSS-VARIABLE INTERACTIONS**

A variable’s relationship with credit risk may depend on other variables within the dataset. For example, the effect of being in a geography or market segment, or the effect of negative public sentiment on a company’s probability of default may be dependent on the size of the company. With smaller companies, there is often less public sentiment data to develop a robust signal.

**Exhibit 1: Example: Non-linear risk signal for a company profit margin**

While profit margin is negative, there is a clear inverse relationship with PD. As profit margin crosses zero, there is a significant downward level shift in risk. Risk is relatively flat for companies with positive profitability.

Source: Oliver Wyman analysis
Exhibit 2 illustrates the impact of variable interactions on the likelihood of default. Four quartiles of an independent variable (Variable 1) have the same density of “bads” as “goods,” for a constant probability of default (PD) of 27.5%. Due to the constant density in the data, a traditional regression model would not identify this variable as a useful predictor. On the other hand, capturing the interaction between the two variables identifies two powerful relationships that are otherwise obscured. We see that Variable 1 interacts with a second variable (Variable 2). In bucket 1 of Variable 2, the quartiles of Variable 1 have an increasing relationship with default (PD increases from 5% to 50%). However, in bucket 2 of Variable 2, the Variable 1 relationship is decreasing (PD decreases from 50% to 5%).

Traditional regression models might not inherently capture these non-linearities. However, it is possible to develop regression models that capture non-linear effects through explicit segmentation, as well as by manually specifying variables through transformations and/or interaction terms (products of two or more variables).

Advanced tree-based algorithms, such as random forests, capture both types of non-linear effects. By using embedded “if” statements, these algorithms effectively slice the data into regions that fit the data most closely. On the other hand, we have also observed that thoughtfully tuned, regularized regression models can approximate the performance of advanced statistical learning techniques and capture non-linearities in one variable or interaction terms. These are clearly not distant in format to traditional regression models, and thereby much easier to explain and implement.

---

**Exhibit 2: Variable 1’s relationship to default as it interacts with Variable 2**

| Bucket 1: | Bads: 110  
| Goods: 290 | PD = 27.5 |
| Bucket 2: | Bads: 110  
| Goods: 290 | PD = 27.5 |

<table>
<thead>
<tr>
<th>Bucket 1</th>
<th>Quantile 1</th>
</tr>
</thead>
</table>
| Bads: 5  
| Goods: 95 | PD = 5% |
| Bucket 1 | Quantile 2 |
| Bads: 20  
| Goods: 80 | PD = 20% |
| Bucket 1 | Quantile 3 |
| Bads: 35  
| Goods: 65 | PD = 35% |
| Bucket 1 | Quantile 4 |
| Bads: 50  
| Goods: 50 | PD = 50% |

<table>
<thead>
<tr>
<th>Bucket 2</th>
<th>Quantile 1</th>
</tr>
</thead>
</table>
| Bads: 55  
| Goods: 145 | PD = 27.5% |
| Bucket 2 | Quantile 2 |
| Bads: 55  
| Goods: 145 | PD = 27.5% |
| Bucket 2 | Quantile 3 |
| Bads: 55  
| Goods: 145 | PD = 27.5% |
| Bucket 2 | Quantile 4 |
| Bads: 55  
| Goods: 145 | PD = 27.5% |

Source: Oliver Wyman analysis
INCLUDE A LARGER NUMBER OF VARIABLES

In addition to capturing non-linearities, advanced statistical techniques allow us to use a larger number of variables than traditional regression models, therefore capturing additional dimensions of statistical signal. In a typical commercial scorecard, the number of useful variables would typically not exceed 10, to eliminate potential multi-collinearity issues.

Advanced statistical techniques can better handle a greater number of variables. For example, regularized regressions and decision-tree-based techniques can more efficiently identify important variables (even marginally valuable ones), from a larger list of candidate variables (including nonlinear and interaction terms), and can retain more variables with less risk of overfitting.4

The added flexibility of advanced statistical learning approaches enables significant increases in credit model discriminatory power. Exhibit 3 is one example of performance gains attained on a representative modeling dataset.

PERFORM BETTER EVEN WITH “DIRTY” DATASETS

Compared to regressions, models built using advanced statistical learning techniques suffer less performance degradation as a result of local data pollution, such as missing values.

**Exhibit 3: Model performance gains from capturing non-linearity using advanced learning techniques (out-of-sample Gini score)**

Starting with a traditional regression model as baseline, we observe noticeable performance improvement by introducing interaction terms, and an even higher performance ceiling when using tree ensemble techniques such as Random Forest, or bias reducing techniques such as Gradient Boosting.

![Graph showing model performance gains](chart.png)

**Source:** Oliver Wyman analysis

---

4 These techniques still require rigorous calibration, tuning and performance testing. An improperly tuned model can substantially underperform more basic regression approaches.
or slight changes in the variables’ definition over time. The ability to capture interactions among a broader set of variables allows models to automatically adjust for clusters of noise in the data.

In one client modeling dataset, we tested a combination of traditional logistic regression and tree ensemble models, first on a relatively “dirty” dataset (with very limited imputation, outlier treatments, or other treatments for noisy data), and then on the “same” dataset with much more cleaning applied. The regression models that were trained on the initial “dirty” data, predictably suffered a significant Gini degradation as compared to the “clean” data. On the other hand, the tree ensemble models were more resistant against pollution in the data and experienced a much smaller deterioration in model fit when trained on the “dirty” data relative to on the “clean.”

MAKE USE OF VARIED NEW DATA

Advanced algorithms can train models using a large number of variables that go beyond squeezing incremental power from existing datasets—enabling statistical models to process non-traditional, unstructured content sets, such as natural language and satellite imagery.

UNLOCK THE POWER OF NATURAL LANGUAGE PROCESSING

Models that focus on natural language processing translate unstructured word information into formats that can be more effectively analyzed and ultimately used in decision making. Unstructured data is represented mathematically as thousands or millions of distinct variables, and advanced learning algorithms then train the models based on patterns in these mathematical representations. Advanced learning techniques are necessary to unlock the value of these content sets—since the complexity of these datasets cannot be captured using traditional regression modeling techniques and cannot easily be represented by analysts in a more structured form.

IDENTIFY KEY SENTIMENTS

The sentiment expressed in company announcements is one example of alternative content with a demonstrated relationship to credit risk. Sentiment includes good news (such as the regulatory approval of a new product, or the termination of an enforcement action), bad news (such as a major litigation announcement), as well as neutral information (such as routine announcements about administrative matters). We have used advanced learning techniques to analyze the wording in each announcement and build models that classify individual announcements according to their predicted sentiment. The calibration followed a smaller dataset where credit officers read the announcements and assessed the sentiment. Exhibit 4 illustrates such sentiment analysis at a high level.

For natural language processing applications such as this one, deep neural networks have been shown as a particularly effective learning technique, due to neural networks’ ability to represent the context of words in many dimensions.
Exhibit 4: Overview of sentiment analysis

**NATURAL LANGUAGE**

ANNOUNCEMENT 1:

ANNOUNCEMENT 1:

**ANALYTICAL ENGINE**

**SENTIMENT IDENTIFICATION**

ANNOUNCEMENT 1: Positive

ANNOUNCEMENT 1: Negative

Source: Oliver Wyman analysis

Exhibit 5: Example: A company’s monthly credit-related sentiments

In Exhibit 5, the illustration shows the credit-related sentiment for a company on a monthly basis, tracks aggregate sentiment over time, and allows users to zoom into the contributing factors for each monthly indicator. As seen, it is dynamic and updated with the arrival of new information.

Source: Oliver Wyman analysis

Copyright © 2018 Oliver Wyman
INCORPORATE DATA THAT IS UPDATED IN REAL TIME

Traditional commercial credit scorecards suffer from “stale” risk assessments. Financials used as inputs are only periodically updated and considered a backward-looking metric. Recent information is often brought in using qualitative factors that may be subjective or limited in breadth.

Today, additional content sets are updated frequently and increasingly accessible. While traditional scorecards can be updated as soon as compatible data becomes available, many new data sources come in widely different and often incompatible formats—making it complex to rapidly process, incorporate, and interpret the data. With advanced statistical learning models, predictions are continually refreshed using a wealth of real-time data, such as natural language sentiment data—allowing for timely risk analysis.

By measuring the impact of social media, announcements, and news sentiments to future observed changes in credit risk, it is possible to overlay adjustments on financial-based predictions that incorporate forward-looking, new perspectives and most closely reflect the actual state of the world. Moreover, with the advances in large scale computing, updates even for large datasets can be done in real time.

However, while these advanced techniques have significant benefits, they come with certain challenges and potential pitfalls that may test a bank’s cultural readiness.

FINANCIAL INSTITUTIONS NEED TO:

**Foster business engagement:** The education and enrollment of senior stakeholders, particularly experienced credit officers, are essential for future investment and implementation.

**Leverage technology and employee talent.** Implementing a large-scale computing infrastructure and employing and training talent with the necessary coding skills is essential.

In the next section, we discuss the cultural, regulatory, governance and infrastructure challenges leading banks are facing to adapt advanced models.
CHALLENGES AND POTENTIAL PITFALLS

Hurdles that institutions are wrestling with today when developing and implementing advanced techniques

While the statistical learning techniques described above have significant benefits, they are not a silver bullet. There are significant caveats that need to be kept in mind when using advanced techniques to support commercial credit risk analytics. We highlight five of these below.

1. Good data – a must have
2. Insights dependent on available data signals
3. Complexity and opaqueness
4. Risk of overfitting
5. Non-trivial validation, governance and computing

GOOD DATA – A MUST HAVE

The ultimate quality of predictive modeling depends on the data collected, cleaned, indexed, and adjusted. While this is obvious for any statistical endeavor, we want to emphasize the amplified importance of performing exploratory data analysis, devoting serious time and effort to sanity-checking the data, and exploring unexpected features in statistical learning.

Without understanding the data, auto-piloted data science will be misguided and can easily land you somewhere other than the intended destination. With the vast volume of unstructured data such as news articles and social media posts, exploring and cleaning the data can seem like an unfathomable task. The good news is that, with the right expertise, advanced learning techniques can be leveraged to gain a better understanding of very large quantities of unstructured data.

In addition to the quality of data, modelers must have access to sufficient quantities of data. For example, a consideration in corporate credit is to focus on credit migration as a dependent variable (given its relative richness and applicability) rather than default (which, particularly in corporate credit is likely to be sparse, and may be driven by idiosyncratic factors that are not broadly applicable to making predictions across a population of obligors).
INSIGHTS DEPENDENT ON AVAILABLE DATA SIGNALS

All statistical learning techniques are ultimately dependent on data. Therefore, for any given training dataset, the quality of the outcome depends on cleanliness of the data. The model’s predictive power is constrained regardless of the technique used to train a model. Using advanced algorithms can provide better performance than traditional regression models, but the improvement ceiling depends on how much hidden information is still contained in the data.

We have tested advanced learning techniques on datasets where Gini improvement was up to 15 to 20 percentage points, and on other datasets where there was no incremental improvement—suggesting that existing techniques exhausted the data, leaving nothing to be gained by learning from non-linearities.

Under such situations, one must turn to alternative content data to seek signals that offer information that is complementary, not merely overlapping.

COMPLEXITY AND OPAQUENESS

Statistical learning techniques are often studied and well-understood by leading academics and technical practitioners. Furthermore, these techniques have migrated from the research community to off-the-shelf subroutines in multiple programming languages. However, they are more complex than their traditional counterparts, and thus for the non-technical practitioner, somewhat difficult to penetrate the workings.

For example:

1. It is less straightforward to assign weights to each variable (as in the case of coefficients).
2. With ensemble methods, it is hard to visualize “the whole equation.”
3. Some learning methods are non-deterministic.

It is important to note that the complexity and opacity do not mean that methods are a “black box.” The specific quantitative mechanisms are well-understood. However, the degrees of flexibility in these models mean that coming to a solution becomes more of an empirical question of performance and stability—a less “closed form” solution than credit practitioners would typically be used to.

In any case, these models come with a host of assumptions, limitations and weaknesses which need to be studied in-depth and must go through comprehensive developer testing, sensitivity analysis, data controls, process controls, and independent validation.

It is essential to consider business challenges that may be introduced by advanced learning techniques. For example, real-time updates to credit ratings have a risk of being volatile over
time. Precision also needs to be balanced against relative stability and predictability. And, while the models may be well understood by modelers and validators, they may present challenges with respect to communicating with customers—such as the non-monotonic and non-continuous outcomes that may be introduced by ensemble techniques.

**RISK OF OVERFITTING**

The flexibility of advanced techniques is a potential double-edged sword. Because they can capture conditions and non-linearities within the training dataset so well, there is a significant risk of overfitting a model to the data. By being disciplined and smart, this can be managed with the following methods:

1. Cross-validation as well as totally out-of-sample testing
2. Ensembles factoring in different techniques
3. Model performance monitoring

**NON-TRIVIAL VALIDATION, GOVERNANCE AND COMPUTING**

Given the recent regulatory focus on model risk management, banks have made significant strides in improving modeling infrastructure and controls. However, the sheer size of many alternative content sets and the high computational requirements of tuning advanced analytics models require developing and implementing sophisticated governance practices.

In regulated environments, strict end-to-end risk management and governance practices around these applications are at the inception and require banks to create development and validation standards for this new breed of models. Modelers may also be challenged with minimizing the biases in these types of models.

Although infrastructure is not the subject of this paper, it is important to note that there are numerous associated engineering and skillset challenges in implementing advanced learning algorithms.
WHERE TO START: LEARNING FROM STATISTICAL LEARNING

LEADING-EDGE USERS – READY TO MAKE THE MOVE

If your institution is ready, we recommend that you deeply engage with your teams, systematically vary the features of these new models, and study the calibration data, including the embedded signals and biases.

Key areas to explore include:

- Theoretical and practical underpinning of the methodologies
- The impact of different out-of-sample testing and cross-validation approaches
- Rigorous back testing
- Understanding granular sensitivities, scenario results and behavior models for outliers and extreme conditions
- Limit behavior

We suggest that you communicate these learnings to all stakeholders—and go through independent model review and validation—before taking the bold step to implement these models.

IF YOUR INSTITUTION HASN’T STARTED YET – HERE’S HOW TO ADAPT

If your organization is not ready to move full speed ahead we recommend a pragmatic “middle path” approach—where you can use statistical learning to guide the institution towards intelligent changes to traditional techniques without abandoning them.

The new techniques can point to what is important, and serves at minimum as a “challenger” approach or provides a more timely monitoring methodology. Pursuing this middle path is a much easier sell to skeptical constituencies and success in this area can lay the groundwork for a fuller conversion to modern techniques.

We have seen banks hesitate for three main reasons: lack of cultural readiness, insufficient model validation, and the cost of implementation. Nonetheless, we believe that institutions
can still leverage insights from statistical learning to improve their existing capabilities. The following section describes practical ways to make use of, and learn from, advanced learning techniques within the context of existing capabilities to:

- Enhance existing models and capabilities
- Develop use cases that are a step removed from decision making
- Bring skeptical stakeholders on board

ENHANCE EXISTING MODELS AND CAPABILITIES

A practical next step for initial adaption is a quick implementation of select, widely agreed-upon learnings, especially the ones that fit within the existing modeling format and computing infrastructure. We have observed that through better variable selection and tuning, institutions can use advanced learning techniques to improve the performance of traditional regression-based credit models. For example, adding select interaction terms identified through regularized regression to expand existing logistic regressions. While this is clearly not the end state, we believe that the realized gains will pave the way for future investments and adoption of more advanced approaches over time.

Several approaches can be used to enhance existing models.

Institutions can:

**Identify variable interactions and transformations** either by examining the structure of more flexible approaches, or via more effective variable selection techniques such as regularized regression.

**Apply advanced approaches** to identify the remaining signal in the residuals or prediction errors of traditional approaches. This information may help detect additional explanatory variables for the model.

BRIDGE THE GAP BETWEEN ADVANCED AND TRADITIONAL APPROACHES

At a high level, these approaches partially bridge the gap between the flexibility of advanced learning techniques and the inflexibility of regression models. A traditional linear model captures statistical relationships in a “flat” manner, without considering interactions, while an advanced learning algorithm captures a great deal of variation and non-linearity in different areas. A regression model with interaction terms curves the prediction space of a traditional model to arrive at a middle ground.

In Exhibit 6, we show a tangible improvement to the accuracy of a probability of default (PD) model. Through using regularized regression techniques, we identified interaction variables that were ultimately included in a traditional logistic regression. Starting with approximately 1,000 potential interactions, we shortlisted 50 candidate interaction variables using
regularized regressions, and identified three incremental variables to optimize the model fit and obtain a 5 percentage point improvement in the out-of-sample Gini.

**Explanation of interaction effects:**
- Impact of Variable A is boosted if Variable B is high as well.
- Variable C is less of a risk indicator for firms with a high Variable D.
- For firms that show high Variable C, the impact of Variable E is much higher.

**IMPROVE MODEL PERFORMANCE WITHIN PARAMETERS OF TRADITIONAL APPROACHES**

The focus for these refinement techniques is to implement models that the institution and stakeholders are comfortable with, and to identify marginally valuable information that squeezes additional performance from the data. For each marginal improvement, it is important to show evidence and explain to the relevant stakeholders why the improvement was possible.

**Exhibit 6: Example – Improving the accuracy of a probability of default (PD) model**

*By using statistical learning techniques, we identified incremental interaction variables to explain the Gini improvement*

**OUT-OF-SAMPLE GINI**

<table>
<thead>
<tr>
<th></th>
<th>0.56</th>
<th>0.58</th>
<th>0.60</th>
<th>0.62</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable A x Variable B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable C x Variable D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable C x Variable E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Explanation of interaction effects:**
- Impact of Variable A is boosted if Variable B is high as well.
- Variable C is less of a risk indicator for firms with a high Variable D.
- For firms that show high Variable C, the impact of Variable E is much higher.

*Source: Oliver Wyman analysis*
By carefully identifying incremental explanatory variables, we have observed that it is possible to improve performance while remaining within the constraints imposed by traditional regression techniques. As a result, these improvements can be implemented within the institution’s existing computing infrastructure, and require minimal updates to documentation or additional validation.

DEVELOP USE CASES THAT ARE A STEP REMOVED FROM DECISION MAKING

To implement advanced approaches, institutions can create use cases for challenger models and borrower early warning indicators—that are a step removed from decision making. This provides the opportunity to develop capabilities that drive incremental value and automation, without taking on the risk of replacing existing proven techniques.

ONGOING MONITORING AND CHALLENGER MODELS

In the short term, we expect traditional models to continue being used for decision making. However, deploying side-by-side advanced learning models provides an effective benchmark, allows analysts to learn from the distinct point of view, and provides useful signals—in case the model outputs begin to diverge rapidly.

Take for example a traditional regression model and a decision-tree-based ensemble that had overlapping top factors at the time of development. By continually re-fitting the tree-based...
model with recent data, it’s possible to use it as a monitoring tool for the traditional model. If the factor contribution ordering in the challenger begins to shift, that is a flag to reexamine and potentially redevelop the traditional model.

In the long run, institutions may “flip” this relationship, using “ready for primetime” advanced learning models for core decision making, while leveraging simpler techniques for ongoing monitoring and challenge models.

EARLY WARNING SIGNALS AND ATTRIBUTION APPROACHES

Approaches that rely on advanced approaches can be effective tools to initiate warning signals for human analysts and help them gain a better understanding of attribution models. In banking, sentiment and transaction analysis models can be used to develop timely warning signals and flag the need for credit analysis review—even if the advanced models are not directly integrated into credit decisions.

For example, a bank uses social media natural language data to develop early warning indicator models that capture both systematic (such as country-level economic sentiment) and idiosyncratic (such as borrower-level financial distress) factors.

We have also seen sentiment models used for “attribution” to highlight the text that is most significant for the overall sentiment flag. Exhibit 7 illustrates one way of displaying such attribution—by using a word cloud of key positive and negative contributing words.

BRING SKEPTICAL STAKEHOLDERS ON BOARD

The challenges arising from advanced statistical learning models are significant for analytical team members and apply even more to other stakeholders. It’s beneficial for CEOs and institutional stakeholders to develop a solid understanding of advanced statistical learning models—to easily converse with direct reports and ask probing questions before problems arise.

Given the complexity and opacity of advanced techniques, coupled with unfamiliarity that we have observed among front office business professionals, validators and regulators have express skepticism. In some instances, intuitive communication and repeated knowledge transfer sessions are needed since the features are discovered in a black box, but the strategy is developed in a white box.

The learnings we discussed are important to develop buy-in among a broader stakeholder group. For front office professionals, warning signals and attribution are useful tools to show the alignment between learning techniques and business intuition.

Using challenger models, as well as carefully tailored and explained refinements to traditional models, is an effective way to give independent validators and regulators familiarity with alternative techniques.
CONCLUSIONS

Creating better, faster and scalable credit risk measures to support commercial credit decisions for underwriting and dynamic portfolio monitoring has become a reality at leading banks. However, cultural, regulatory and infrastructure driven challenges are holding banks back from further accelerating the development and implementation of a new breed of credit risk models. The benefits are however significant, and it is a matter of time before these models are used widely.

If your bank is ready, we encourage bold steps. We recommend that you deeply engage your teams with the workings and features of these new models. Some key areas to explore include theoretical underpinnings, sensitivities, outlier analysis, cross-validation approaches. Communication to stakeholders and independent validation are key steps that should be followed thoroughly.

If your bank is hesitating, we recommend a pragmatic middle path—experimenting and implementing what you learned from your experiments in your current setting as initial steps. This directs developers to understand the techniques and data profoundly to explain marginal benefits of each step, while comparing existing models with what these new approaches offer. These learnings from statistical learning will pave the way for future investments and adoption of more advanced approaches over time—leading to better and faster credit risk management.
ABOUT THE AUTHORS

UGUR KOYLUOGLU
Vice Chairman and Partner for Financial Services Americas
ugur.koyluoglu@oliverwyman.com

ATILIO MEUCCI
Partner for Finance & Risk and Public Policy Practices
attilio.meucci@oliverwyman.com

GOKCE OZCAN
Partner for Finance & Risk and Public Policy Practices
gokce.ozcan@oliverwyman.com

SIMON SCHWENDNER
Associate in the Finance & Risk and Public Policy Practices
simon.schwendner@oliverwyman.com

KIRILL SKOK
Principal in the Finance & Risk and Public Policy Practices
kirill.skok@oliverwyman.com

DANIEL WANG
Principal in the Finance & Risk and Public Policy Practices
daniel.wang@oliverwyman.com
Oliver Wyman is a global leader in management consulting that combines deep industry knowledge with specialized expertise in strategy, operations, risk management, and organization transformation.

For more information please contact the marketing department by email at info-FS@oliverwyman.com or by phone at one of the following locations:

AMERICAS  
+1 212 541 8100

EMEA  
+44 20 7333 8333

ASIA PACIFIC  
+65 6510 9700

www.oliverwyman.com