

Fintech Takes



Can LLMs Manufacture Better Credit Decisions?



Taktile

One of the challenges in writing about fintech is that it's not always intuitive how a new technology will reshape the financial services industry.

In other industries, it's more obvious because the mechanics of how those industries work are more concrete. They have tangible inputs and outputs and relatively straightforward manufacturing processes.

Financial services is different. It's abstract; just a bunch of numbers in different databases constantly being updated by very old computer systems that no one wants to touch for fear of breaking them.

So when a new and potentially disruptive technology comes along, it can be very difficult to figure out exactly how it will impact the financial services ecosystem, even if it's obvious that it will in a profound way.

A good example is the intersection of credit decisioning and artificial intelligence.

We know that AI has already revolutionized the process of determining who to give a loan to. And it's plainly evident that large language models (LLMs) and the transformer architectures they are built on will push that revolution even further forward.

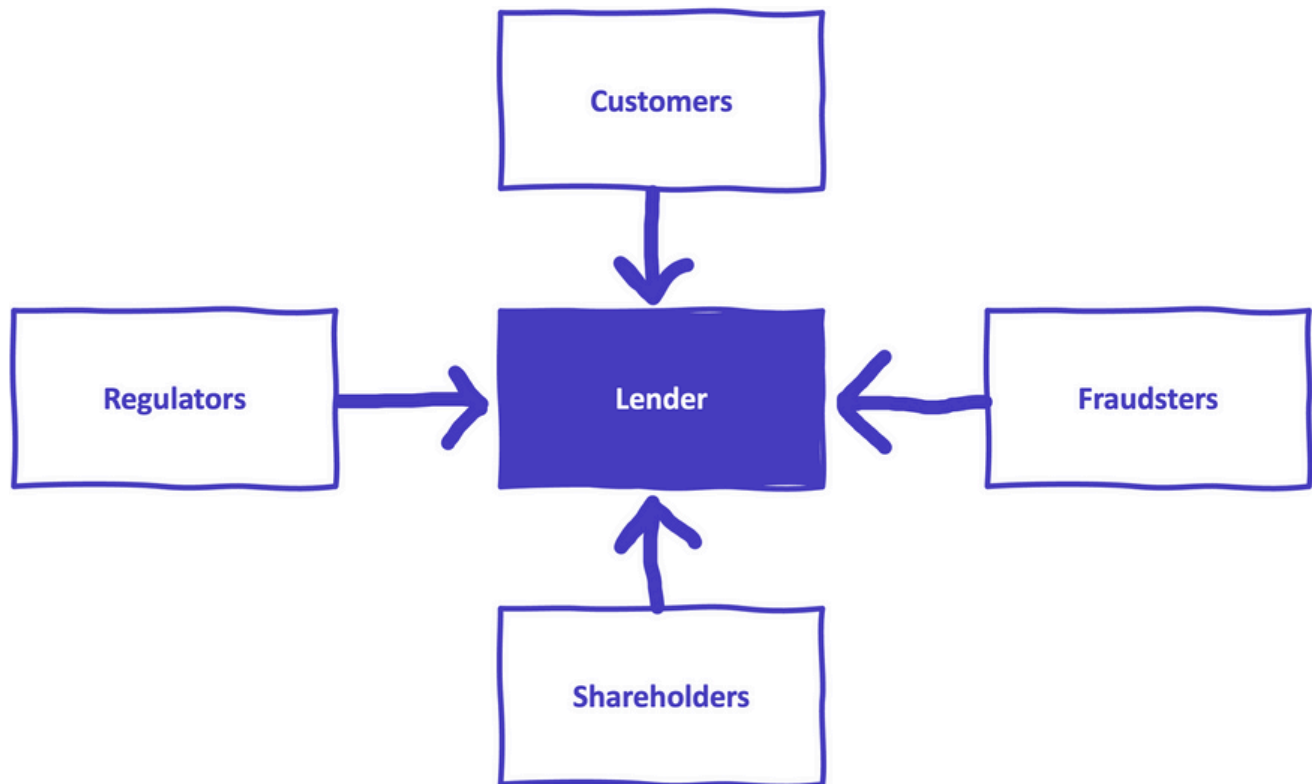
But how, exactly? And how far?

These are tricky questions to answer. To do so, we must first understand how lenders "manufacture" credit decisions. What process do they go through? Why does that process work the way that it does? How has it changed over the last 60 years? And where might LLMs add differentiated value to a manufacturing process that is already extremely well-optimized?

Those are the questions that we will address in today's essay.

The Credit Decisioning Ecosystem

Lenders sit in the middle of an ecosystem, which is populated by four primary stakeholders, all of whom want different things.



Customers want loans, obviously. Ideally, they would like loans that are inexpensive, convenient to acquire, and tailored to their exact needs and circumstances.

Shareholders want profit (again, very obvious). Ideally, they would like risk-free profit, but that is rarely possible in lending, so they are usually willing to accept a reasonable level of risk in order to get the returns they are looking for.

Fraudsters want loans, but unlike legitimate customers, fraudsters have no intention of repaying them. Ideally, lenders would like to completely prevent fraudsters from acquiring loans, but that goal is generally not compatible with shareholders' desire for risk-adjusted returns. Thus, for most lenders, the optimal amount of fraud is small, but not zero.

Regulators want customers to be treated fairly by lenders (in accordance with established laws and regulations), and they want lenders to avoid financing terrorism or other sanctioned activities.

In a perfect world, lenders would be able to maximize value to customers, shareholders, and regulators, while completely stopping any value from accruing to fraudsters.

Unfortunately, we do not live in a perfect world.

Credit decisioning requires tradeoffs, and the process that lenders use to manufacture credit decisions can tell you a lot about how they think about those tradeoffs, especially when a new and disruptive technology is introduced to the market.

There are two examples here that I find instructive:

1. Risk-based pricing. In the 1980s, credit cards were a one-size-fits-all business – most issuers offered the same interest rate and annual fee to everyone that they deemed an acceptable credit risk. Then, in 1994, Capital One appeared. It flipped this model on its head by using data and statistical analysis to segment customers into granular risk profiles and offer customized credit card terms – higher rates for riskier borrowers, lower rates for safer ones. This innovation to the credit decision manufacturing process was incredibly disruptive. It (arguably) benefited subprime customers, who were more able to access mainstream credit. And it (inarguably) benefited Capital One's shareholders, who were thrilled that the bank had found a profitable way to grow in an already crowded market. However, it made regulators nervous. They liked the increased access to credit for subprime borrowers, but they worried about the potential of data-driven pricing and marketing to empower more predatory lending models.

2. Online account opening. As the internet became a more popular distribution model for lending products in the late 1990s and early 2000s, lenders were forced to adapt their credit decisioning processes to keep pace with their customers' expectations. The paper forms and manual underwriting processes that defined branch-based account opening processes were slowly replaced by digital applications and automated, real-time underwriting systems. This evolution was appreciated by customers (who value convenience above all else) and shareholders (who tend to like anything that drives growth), but it was also a gift to fraudsters (who found online account opening to be much more efficient and low-risk than branch-based account opening).

There was no way that data-driven customer segmentation and digital distribution channels weren't going to disrupt the lending industry. If it hadn't been Capital One and early digital lenders like Bill Me Later and LendingClub, it would have been others.

But it's important to remember that these disruptive innovations weren't immediately or universally successful. It took a long time for them to become ubiquitous, and neither of them solved all the problems that their early proponents thought they would solve.

The reason for this is simple, and important to remember – **there is no technology, no matter how innovative, that is a silver bullet in credit decisioning.**

The best we can hope for is to incorporate these new technologies thoughtfully, in a way that helps lenders move outcomes for a specific group of stakeholders in the right direction, without requiring a significant tradeoff from the others.

So, in order to figure out how LLMs can be applied, thoughtfully, to credit decisioning, we need to examine how the process of manufacturing a credit decision works.

Manufacturing a Credit Decision

To dramatically oversimplify, every credit decisioning process is built on three fundamental components: **data** (what you know about the customer), **predictions** (data-driven guesses about how the customer will behave in the future), and **rules** (a logical set of steps for using customer data and predictions to make decisions).

These components are the building blocks for the infrastructure that is used to address three essential questions, the answers to which are enormously important to the various stakeholders outlined above:

- **Is the customer who they say they are?** (This is where we try to screen out the fraudsters.)
- **Is the lender legally allowed to do business with the customer?** (This is where we ensure that we are complying with regulators' anti-money laundering rules.)
- **Does the lender want to do business with the customer, and if so, at what terms?** (This is where we hopefully find a balance between shareholders' interest in generating risk-adjusted revenue and customers' interest in getting a loan, while also ensuring that regulators' fair lending rules are complied with.)

Historically, the way we answered these questions was highly manual. A customer would walk into a bank branch and fill out an application for a loan. A loan officer would call up a credit bureau and ask someone there to read the relevant portions of the customer's file to them over the phone. The loan officer would then evaluate the customer (using a combination of objective and subjective criteria) and decide whether to give them a loan or not, and, if yes, at what price.

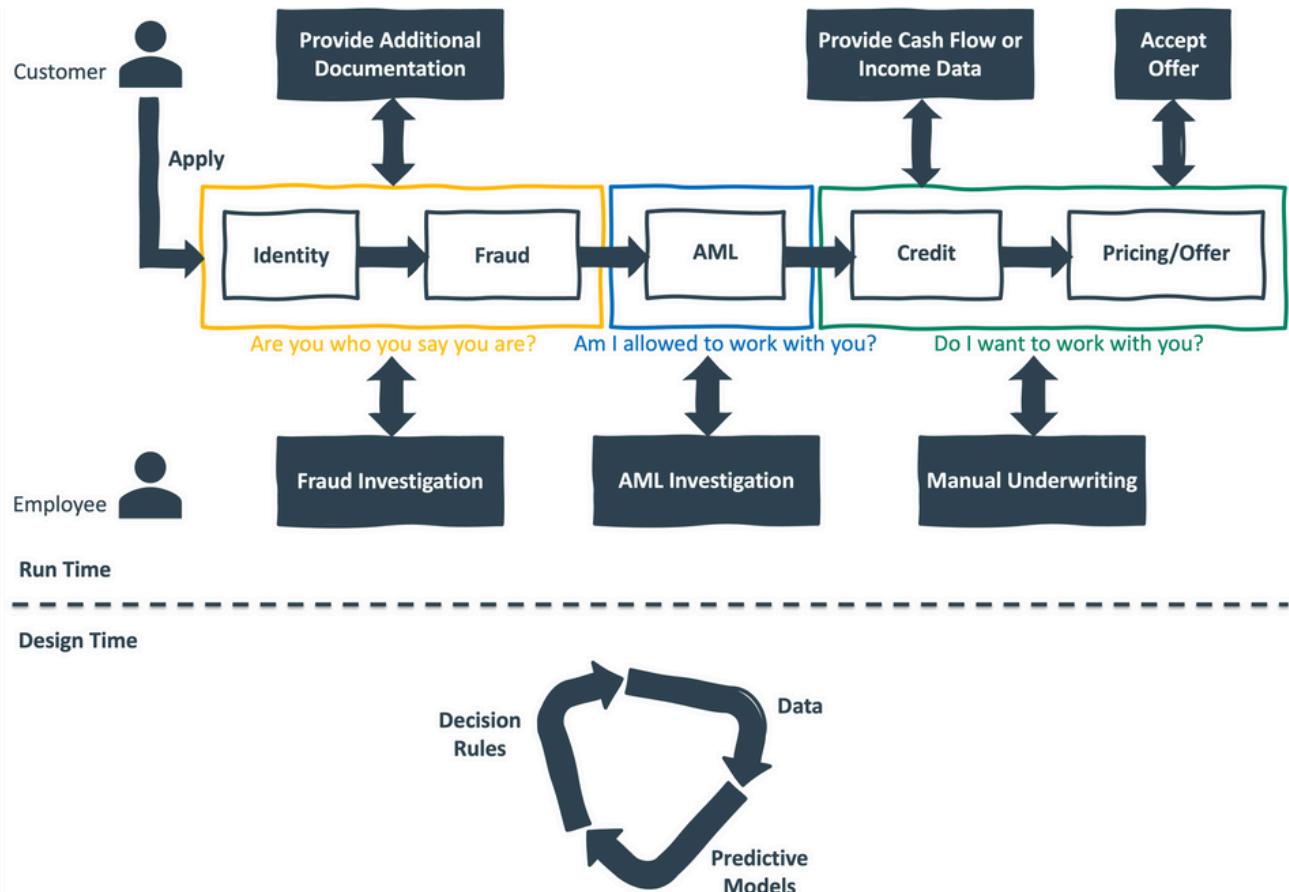
This manufacturing process was slow, expensive, and riddled with errors and bias.

However, thanks to the thoughtful application of technology, it has become significantly faster, cheaper, and more accurate and fair over the last 60 years. All three of the fundamental components of credit decisioning have gotten big upgrades during that time:

- **Data** — Between 1960 and 1990, the credit bureau industry consolidated down to the big three that we have today, and those companies moved from paper-based record systems to electronic databases. This made credit data far more fair and accurate (since it was shared system-to-system rather than manually) and much easier for lenders to utilize (credit data became far more structured and standardized in the 1980s).
- **Predictions** — The first big step forward here was FICO, which started by helping individual lenders apply statistical analysis techniques to predict customers' likelihood to repay a loan in the 1950s. As the credit bureau system became more standardized and digitized in the 1980s and 1990s, FICO introduced general-purpose predictive scoring models, which could be used, off the shelf, by any lender. In the 2000s and 2010s, lenders' predictive powers took another leap forward thanks to advancements in computing power and data analytics, which led to the development of more sophisticated statistical models (built using techniques like logistic regression) and machine learning (ML) models (using techniques like decision trees, random forests, and gradient boosting machines).
- **Rules** — Starting in the late 1980s and early 1990s, lenders began to supplement human loan officers with software. These rule-based credit decisioning systems, which leveraged electronic credit data and predictive scoring models, allowed lenders to codify the best practices of their most successful loan officers, while dramatically increasing the speed and scalability of their credit decisioning processes through automation. More modern SaaS-based credit decisioning systems also allow for greater agility in decision-making through configurable rules that can be adjusted without extensive reprogramming. This enables lenders to design and test highly segmented, real-time decision workflows and models tailored to their customer bases.

The result of all these upgrades is a credit decision manufacturing process that, while not perfect, is very well-optimized to deliver the outcomes that customers, shareholders, and regulators want (and to foil the outcomes that fraudsters want).

Here's an oversimplified view of what that process looks like:



A couple of quick notes:

- Run time (the top two-thirds of the graphic) is the (mostly) automated process happening in real time to evaluate customers' loan applications. The goal is to be as efficient as possible, but (where necessary) the lender will reach back out to the customer to get additional information or documentation. Applications can also be routed to the lenders' employees when there is a suspicion of fraud or money laundering, or when the "Do I want to work with you?" question is too close for the automated system to make the call.

- Design time (the bottom third of the graphic) is the continuous work that the lenders' data scientists, business analysts, and engineers do to refine and improve the lender's automated credit decisioning system. This includes analyzing the lender's performance data and third-party data sources, utilizing machine learning and other analytical techniques to develop or improve attributes and predictive models based on that data, and building and testing new decisioning rules or logic for processing loan applications. This design-time process is represented as a loop because, generally, the goal is for this work to be as rapid and iterative as possible, allowing lenders to be agile in responding to changing market conditions and customer behaviors.
- Decision rules and ML-powered predictive models require structured data to work effectively. As a result, most of the data that is traditionally acquired during the credit decisioning process (credit bureau data, fraud consortium data, etc.) is highly structured, and any data that is collected directly from the customer must be transformed into a structured format for it to be useful (e.g., loan applications for collecting PII, OCR models for scanning documents and extracting data, etc.) Additionally, ML models perform better when they are trained on large datasets that include clear, closed-ended outcomes. Evaluating credit risk, where outcomes are always clear (they paid or didn't) and large historical datasets are easy to acquire (internally or from the credit bureaus), is a good example of a job-to-be-done that can be done very well with machine learning.
- By contrast, whenever a human needs to be pulled back into the credit decisioning loop (e.g., customers being asked to provide additional data, fraud/compliance/credit analysts being assigned a new case to work, etc.), they generally prefer to be communicated with using natural language as opposed to structured or tabular data, which takes longer for them to parse.

It bears repeating that this process is already very well optimized.

It's not perfect, obviously. It still requires a decent amount of human involvement, which is generally bad because humans are expensive, slow, and prone to subtle errors and biases. It's allergic to unstructured data, and, as a result, it tends to work much better for lending products and customer segments where there is an abundance of structured data available to train models and make decisions than it does for products and customers that lack structured data.

But still. It works well.

So, if we want to add large language models into the credit decisioning process – and given the rapid and impressive advances in LLMs that we have seen over such a short time, we probably should – the question we need to ask ourselves is where, specifically, can LLMs add value?

This is the question that we will end today's essay on. And to start us off, we should do a quick refresh on what LLMs are, how they work, and why they are fundamentally different from machine learning.

A Quick Refresh

Traditional ML models, like logistic regression or random forests, are all about structured data and specific predictions. Imagine a random forest predicting loan defaults: it's handed a table – credit scores, income, payment history – and trained to output a probability or a yes/no call. These models shine with clear inputs and defined outcomes (paid or didn't), making them a perfect fit for credit decisioning's standardized datasets. From FICO's statistical roots to Capital One's risk and marketing segments, ML has thrived on the tidy, tabular world of loan servicing systems and credit bureaus.

LLMs are different. They use a mechanism called “attention” to process unstructured data, like emails, call transcripts, or loan applications. Attention weighs every word’s importance relative to every other word in a sequence, all at once. In “The borrower who defaulted last year reapplied today,” an LLM links “borrower” to “defaulted” and “reapplied,” no matter the gaps, capturing context where more traditional analytic models may get tripped up. They’re trained on vast datasets – think billions of pages from books and websites – to predict the next token in a sequence, then fine-tuned for tasks like answering questions or summarizing documents. This makes them less about single-purpose number-crunching and more about understanding inputs and generating outputs.

Traditional ML is task-specific: a decision tree predicts risk, nothing more. LLMs are versatile – capable of classifying risk, drafting loan terms, or parsing customer chats from one model. ML demands structured tables; LLMs excel with unstructured text, like a borrower’s written appeal. ML delivers numbers or labels; LLMs produce contextual, human-like language. In the highly structured world of lending, ML has reigned supreme. But where the process hits snags – human intervention or messy inputs – LLMs can add a lot of value.

And they’re improving all the time! Over the last three years, AI labs have made significant advances in software optimization and model fine-tuning, which (along with constantly improving GPUs and a growing shift towards open source) have significantly reduced the costs to develop and deploy LLMs, thus making them more accessible to a broader range of companies (Andreessen Horowitz has termed this trend LLMflation).

Additionally, recent innovations like Prior Labs’ TabPFM extend LLMs’ underlying transformer architecture into tabular data – ML models’ bread and butter – matching or beating classics like XGBoost on small datasets in a single pass. These leaps mean LLMs aren’t just for unstructured data anymore; they’re encroaching on ML’s turf, and doing so with increasing accessibility and affordability.

So, if you wanted to apply LLMs (and the transformer architecture that underpins them) to improve the performance of specific steps within the credit decision manufacturing process, where should you start?

Great question!

Applying LLMs In Credit Decisioning

The key is to target the cracks – those areas where human involvement slows things down or where unstructured data gums up the works. There are many such cracks in the credit decisioning process where LLMs could add value, but I'll focus on four.

1.) Turbocharging Investigations

For the same reasons that ML works well in credit risk evaluation, it often underperforms in fraud and compliance risk evaluations. This Taktile [blog post](#), written by AI/ML expert [Peter Tegelaar](#), elaborates:

Unlike credit risk, where ground truth exists in the form of repayment history, fraud [and compliance risk] often lack definitive labels and structured datasets. This is where LLMs shine. Their ability to process unstructured text – such as invoices, emails, and news articles – enables them to uncover patterns and inconsistencies that would otherwise require human intuition.

Today, when automated systems flag applications that are suspected of being fraudulent or in violation of regulatory requirements, the applications are sent to human analysts for further investigation. This is an area where small efficiency gains – reducing false positive rates with better entity matching, streamlining the investigation process by surfacing the most relevant insights using natural language, etc. – can really add up (again, humans are expensive!)

2.) Streamlining Customer Interactions

When the automated process hits a snag – say, needing more information to answer the “Do I want to work with you?” question – customers get an automated voice call, email, or text message. Humans prefer natural language, but crafting those messages and interpreting replies takes time and is prone to errors (if you've ever tried to navigate an “intelligent” IVR system, you know what I mean).

LLMs can power customer-facing AI agents, capable of drafting personalized, clear requests (“We noticed your income documentation is missing – could you clarify your employment status?”) and parsing responses to extract structured data (turning “I’ve been freelancing for two years” into a usable attribute).

3.) Identifying Subtle Risks

Predictive models and decision rules are good for spotting known risks, but often it’s the risks you weren’t even thinking about that can hurt you the most. Human intuition is very good at anticipating these unknown risks, but recently, LLMs have made some impressive gains on this front. Here’s Peter Tegelaar again:

2024 marked a significant turning point in this regard. Advancements in reasoning models led to the development of “reasoning engines.” Popularized by Andrew Ng, this approach emphasizes leveraging LLMs for reasoning over ambiguous data rather than relying on them for factual knowledge, which carries risks of hallucination.

This generalized reasoning capability can be tasked with acting like an ombudsman for the credit decisioning process, looking at it through a systemic lens and spotting subtle signs of risk. Here’s Tegelaar one more time:

One unexpected advantage was LLMs’ ability to detect subtle inconsistencies missed by traditional systems and analysts. For example, they uncovered a fraud ring by identifying recurring grammatical errors across invoices from supposedly unrelated entities.

4.) Accelerating Design-Time Experimentation

The design-time loop – where business analysts and data scientists tweak models and develop new rules – is a crucial source of competitive differentiation for lenders, but only if it's fast.

LLMs can accelerate it by acting as co-pilots. Here's how Maximilian Eber, Chief Product and Technology Officer at Taktile, explained this value proposition for LLMs in a recent blog post:

We have found that AI is very helpful in building out and improving decision logic. Imagine a co-pilot that identifies bottlenecks in the decision process where many applicants are rejected, then recommends data sources to help reduce false positives, and thereby lets you raise approval rates. Most importantly, we believe AI can make decision authors' lives easier without taking away control and transparency over how decisions are ultimately made.

There are many pieces in the logic layer that AI can help with. We have discovered that AI is very good at defining test cases, helping you to find gaps in your logic that you might have otherwise missed. Similarly, AI can help you resolve unexpected errors when handling edge cases, or write code for you to capture complex business logic that is difficult to express with low-code building blocks only.

Credit Decisioning + LLMs: It's Time to Get Started

A lender's credit decisioning process reflects the priorities of the stakeholders in the lender's decisions.

When a new technology comes along that can be broadly useful across the ecosystem, lenders have no choice but to adopt it.

Large language models are one such technology.

Fraudsters are already adopting them to make their attacks more efficient and effective, putting the onus on lenders to adopt them in order to develop viable countermeasures.

Shareholders are pushing lenders to adopt them due, primarily, to the potential for significant cost savings.

Regulators aren't pushing lenders to adopt LLMs (that's just not what regulators do), but they are expressing a growing level of curiosity in the power of LLMs to open up greater access to credit for consumers and small businesses (while still acknowledging that there are significant explainability and accountability concerns).

And customers? Customers just want convenient loans at fair prices. And they will adopt any new technology that gets them closer to that outcome.

That now includes LLMs.

The time to start integrating LLMs into your credit decisioning process is now (and there are providers out there like Taktile that can help).

About Sponsored Deep Dives

Sponsored Deep Dives are essays sponsored by a very-carefully-curated list of companies (selected by me), in which I write about topics of mutual interest to me, the sponsoring company, and (most importantly) you, the audience. If you have any questions or feedback on these sponsored deep dives, please DM me on [Twitter](#) or [LinkedIn](#).

Today's Sponsored Deep Dive was brought to you by [Taktile](#).



Taktile is an AI-powered decision platform that empowers risk, credit, and fraud teams to tap into every layer of the AI stack—data, models, and decision logic—to automate and optimize risk strategies across the customer lifecycle. As generative AI becomes mission-critical in financial services, Taktile helps teams harness its potential with built-in guardrails, real-time data, low-code tooling, and a rich marketplace of third-party integrations—enabling faster iteration, less engineering overhead, and better outcomes at lower cost.