E-BOOK

ZEST AI

MACHINE LEARNING 101
FOR CREDIT UNDERWRITING

TABLE OF CONTENTS

ZEST AI

INTRODUCTION

Machine Learning

Believe it or not, artificial intelligence, machine learning theories, and logic systems have been around since the 1940s. And since then, they've become an increasingly prominent part of our daily lives. Whether we realize it or not, artificial intelligence has helped technology make our lives more efficient.

So… what is it?

Machine learning is important in helping us navigate large volumes of data and trends to make better informed decisions — from providing recommendations while online shopping or movie streaming, to taking the best route to work, to fraud detection, and of course credit underwriting.

By the end of this ebook, you'll learn:

- What is machine learning, and how do we use it?
- How does machine learning work, exactly?
- And, most importantly… How does machine learning affect credit underwriting?

Machine learning (often abbreviated as ML) is a subset of the broader field of artificial intelligence (AI), primarily focused on predicting outcomes. When coding or writing a program is too time-consuming, not scalable, insufficient, or impossible due to the amount of information required and its complexity, machine learning can be a valuable tool to accomplish tasks. The computer mimics how humans learn — making observations, testing variables, and making predictions based on cause and effect.

Writing code for software is more like a recipe. You have to know which ingredients to use and how long to cook them to create your desired meal, or know every detail about how to execute a program. If you want to create individual recipes for all the ways to cook meals with every possible ingredient, it's very difficult to scale this task.

A powerful computer can learn and make these predictions using vast amounts of information and can make associations between different variables very quickly. In other words, machine learning can take that vast amount of data (ingredients) and quickly generate all those recipes.

In the next section, let's run through some examples, and discover how machine learning can benefit the financial industry.

ZEST AI

CHAPTER ONE

Practical applications of machine learning

Machine learning is a sophisticated tool to help us predict outcomes. These predictions can be as simple as answering, “What do I buy at the grocery store?” or as complicated as, “Does this credit union member qualify for an auto loan?”

For example, when you shop, you often purchase items together. If you buy pasta, you might also buy pasta sauce. When you go to the grocery store, you’ll find that the boxes of pasta sauce are located next to the jars of pasta sauce. How convenient!

Machine learning employs a similar logic to help people make decisions. For example, when you shop online and buy a pair of snow boots, you might receive an ad for a snowboard. Machines learn how to predict what shoppers will purchase, when, and with other products to make purchasing decisions faster and easier. In addition to the items themselves, the machines will also consider cost, availability, age, etc., and the complicated relationships between these variables to determine your shopping experience.

With credit underwriting, machine learning helps lenders make accurate and consistent predictions about a member’s risk for delinquency. While underwriters can manually review applications, they have limited insights from the variables they can consider, which are usually based on a single score and a few historical applicant details. Not only does this lead to decreased economic performance, but it's slow. And no human is perfect, meaning our judgment can create inconsistent decisions from one application to the next, and perpetuate inherent and unintentional bias.

Using machine learning for credit underwriting comes with many benefits. Not only can machine learning predict how likely a member will be delinquent, it can also:

- Make more accurate risk assessments about members that would traditionally be unscorable with thin files.
- Help debias decisions to ensure equitable lending.
- Provide deeper insights about why a member may need further assistance.

ZEST AI

CHAPTER TWO

Looking under the hood

02

All of this learning and predicting starts with data… a lot of data. All of this information is used to develop a **model**, or a program that can recognize patterns.

To be able to develop a good model with accurate and inclusive predictions, we have to provide enough information for the model to understand relationships and correlations between different **data features**. A data feature is simply a piece of information that can be measured, like household income, or for our purposes, delinquency rates, credit utilization, and so on. A good model will be able to find trends over time and weigh these different features appropriately.

If we look under the machine learning hood, there are a few key parts to building a model:

- **Collecting** and preparing development data sets for training and validation.
- **Training** a machine learning model to correctly learn patterns from the data.
- **Testing** and deploying a machine learning model.

Before you can use a machine learning model to start assessing credit applications, it needs to be trained.

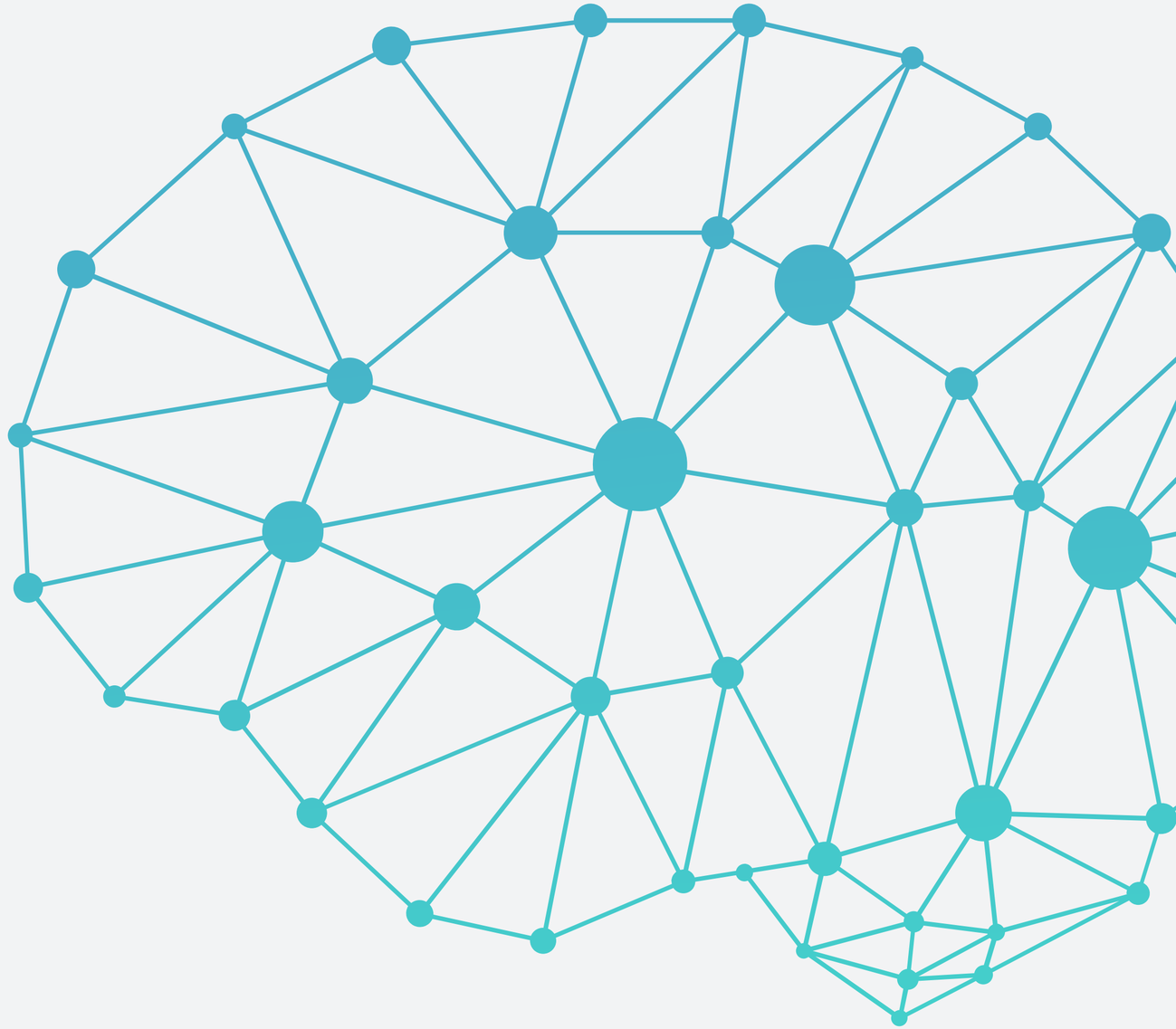
Before it can be trained, data scientists collect and prepare a set of relevant data. They split this data into two sets: one to create and develop the model, and another to test for accuracy. This second set of data is often referred to as a "holdout set."

This data can consist of thousands of different features. For example, Zest AI's technology extrapolates around 10,000 unique pieces of borrower data from years of raw credit bureau reports. You don't need to use 'sneaky' data to gain a better understanding of borrower risk — Zest AI only uses data from our customers, who are regulated by organizations such as the Consumer Financial Protection Bureau.

To train models, we run all of the data through **algorithms**, or a set of instructions that answer a question or solve a problem. Algorithms can help us answer questions like: how likely am I to default on a loan?

CHAPTER TWO: LOOKING UNDER THE HOOD

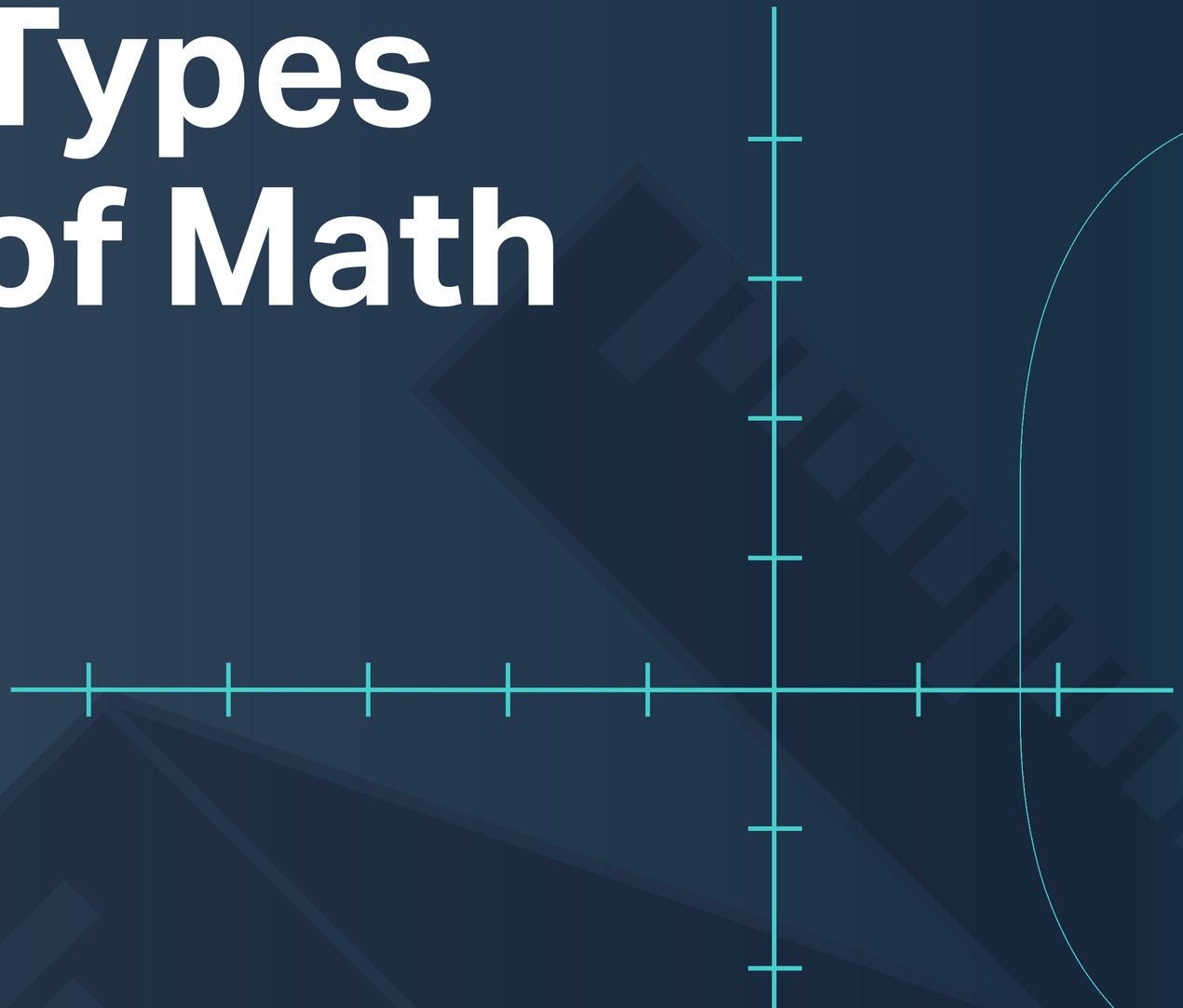
To run data through algorithms, we use different types of math like statistics, calculus, and linear algebra to find out which data, trends, and correlations make the most accurate predictions. Let's take a look at what these different types of math are, and how they can help us solve problems and predict outcomes.



ZEST AI

CHAPTER THREE

Types of Math



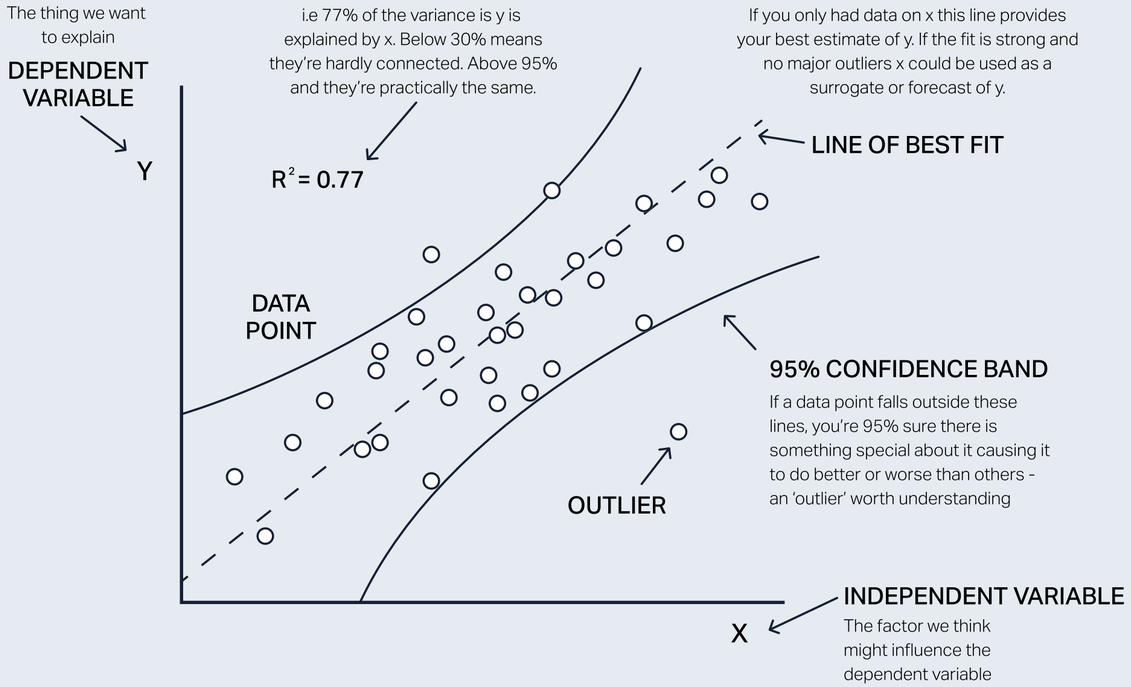
03

Math — specifically the math used in machine learning — tells us about the relationships between different variables. This chapter explains three fundamental tools that allow us to understand complex patterns in their data and make predictions, such as a borrower's payment history and their likelihood to pay their loans.

Let's take a look at some of the different ways we can express these relationships — with numbers!

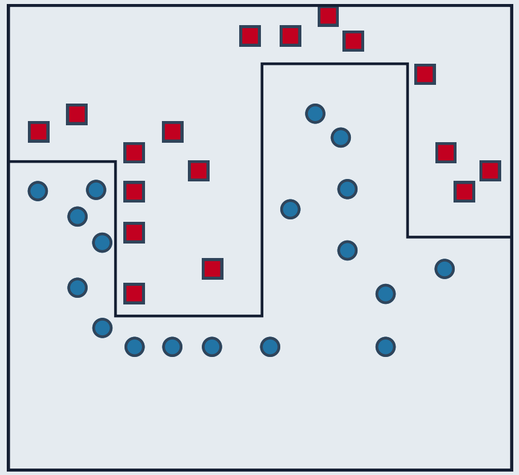
Using **linear regression analysis**, we predict values along a spectrum. Linear regression is a technique that looks at the sum of our predictor variables, like payment history, length of credit, etc., to determine credit risk (e.g., probability of default). You might find that as one's length of credit increases, their risk decreases. The closer the "line of best fit" (shown in the illustration) is to the data points, the more accurately that "line of best fit" is able to predict outcomes.

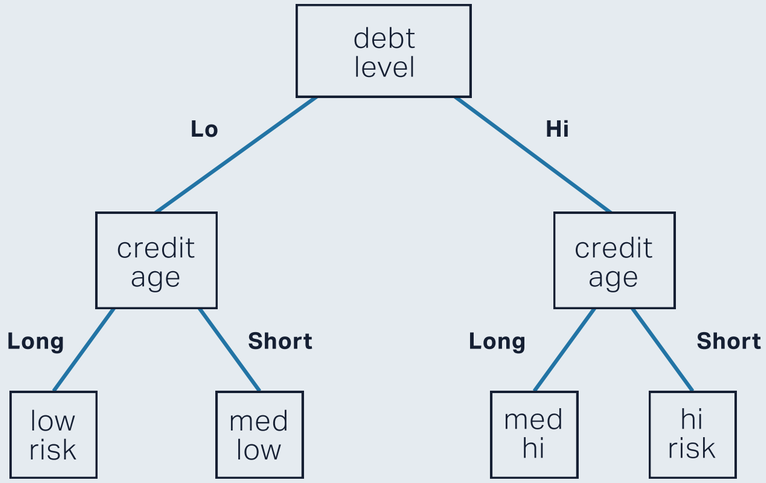
LINEAR REGRESSION



Using regression analysis might not be enough to accurately predict risk. In addition to these types of analyses, machine learning algorithms will use **decision trees** to synthesize a lot of information that often doesn't fit neatly into a straight line.

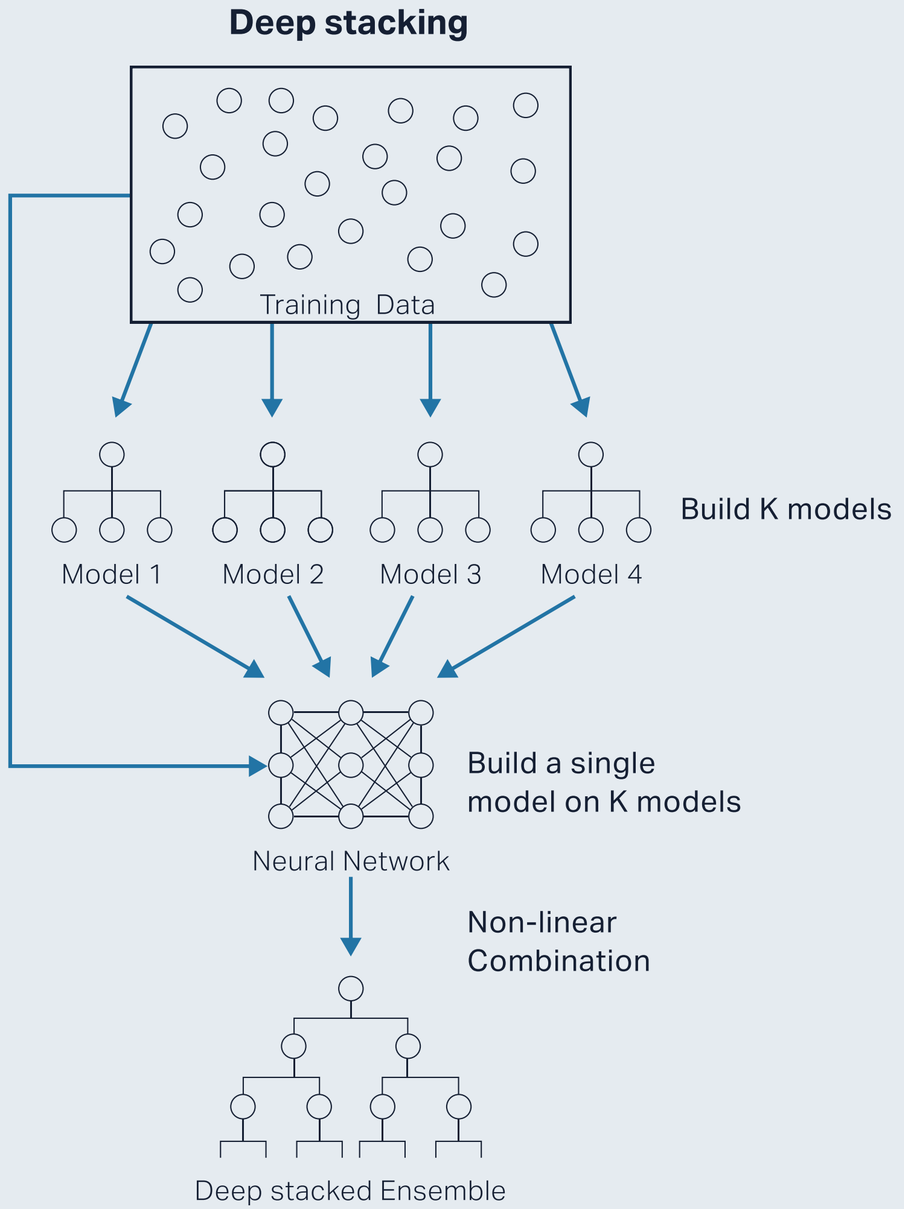
Think of a decision tree as a flow chart. First, we might look at a variable like credit limit — everyone with a credit limit of more than $6,000 will move into one bracket, and less than $6,000 into another. This process will continue for every single data feature, until each borrower is in their correct flow chart spot.





Tree ensembles are many decision trees linked together. A single decision tree can be very precise, but we want to make sure that it is accurate for a wide variety of people. We link decision trees together and look for averages or discrepancies between results to increase our accuracy across a broad range of data.

From thousands of data points, Zest AI develops models with hundreds of features created from these algorithms to accurately determine risk. These methods examine all the deep information within someone's credit profile that determines risk, and will appropriately remove variables that are not proven to gauge credit risk. The additional information realized by the advanced algorithms also allows the model to accurately score traditionally unscorable members such as those who have thin files.



ZEST AI

CHAPTER FOUR

Machine learning in credit underwriting

04

Machine learning is revolutionary for credit underwriting as it solves for many inadequacies of the traditional systems, as long as the models are:

- Explainable
- Accurate and fair
- Constantly monitored

Explainability is critical for credit underwriting. While complex, good machine learning can quantify how much of each variable or feature contributed to a final outcome. Using advanced model explainability techniques, models can be fully documented, all of which is captured in development and production outputs, like a Model Risk Management report and adverse action reasons.

Fairness in lending is one of the biggest gaps machine learning can address. In addition to accuracy, models can be trained to test against disparate impact, so that there is no feature that correlates unfairly with a specific protected class. By using better math and more data with these algorithms, it's possible to meet the dual goal of increased accuracy and fairness.

Zest AI is able to leverage a variety of proprietary and patented techniques to more effectively search for Less Discriminatory Alternatives (LDAs), models that can achieve similar levels of accuracy that result in fairer outcomes for protected class groups.

To ensure appropriate compliance, underwriting models developed with machine learning technology are generally **supervised** machine learning models. This makes them different from the types of “generative” or “unsupervised” AI models that are more commonly in the news. The model is locked down — and never “learns on its own”' or changes on the fly.

Additionally, **regular monitoring** ensures further consistency by monitoring how the different features of the model are distributed, ensuring that the model scores do not drift over time.

Machine learning is an incredible tool that allows us to have better insights to make smarter, more consistent decisions at scale. AI-automated underwriting, like all other machine learning that we use, allows us to do more, faster. It allows lenders to quickly make smarter and fairer decisions, helping members move along their financial journeys.

More data and better math — combined with a commitment to create smart, inclusive, and efficient models — makes AI the future of credit underwriting.

About Zest AI

Faster decisions, expanded access, happier members — Zest AI is a tech company on a mission to make fair and transparent credit accessible for everyone. Since 2009, Zest AI has been innovating and perfecting AI credit underwriting technology. A CUSO since 2021, Zest has over **100 credit union customers**, has partnered with the **credit union leagues**, is running **over 250 active models**, and is helping our customers better **serve over 13 million members** across the country. Zest aims to make best-in-class AI-automated underwriting technology available and accessible to all credit unions.

SMART

Lend confidently and make better credit decisions with deep, accurate lending insights.

INCLUSIVE

Help the underserved and ensure all of your members get a fair shot with inclusive, equitable lending.

EFFICIENT

Deliver more decisions faster with powerful AI, without increasing underwriting resources.

Visit us at zest.ai or contact us at hello@zest.ai to learn more.

GLOSSARY

Term	Definition
Artificial intelligence (AI)	A field of computer science that focuses on building machines that can solve different tasks and problems.
Machine learning (ML)	A subset of artificial intelligence that focuses on using data and algorithms to make accurate predictions.
AI-automated credit underwriting	Using machine learning algorithms to provide an instant lending decision. The machine learning algorithm can automatically and accurately predict credit risk, allowing the lending decision to take seconds.
Model	A program that has been trained to recognize patterns and make predictions based on data.
Data feature	A piece of information that can be measured.
Algorithm	An equation that solves a logical problem.
Linear regression analysis	A mathematical method that shows the relationship between two or more variables.
Decision tree	A model that resembles a flowchart, showing many decisions and their outcomes along a branching pattern.
Tree ensemble	Multiple decision trees linked together that weigh different variables and their outcomes.

ZEST AI

GLOSSARY

Term	Definition
Explainability	The ability to explain a model's results, for instance, why a member was approved or denied a loan. This is especially important when using AI-automated underwriting, as multiple, interrelated factors contribute to the decision.
Fairness	The law dictates that no data feature used in a credit underwriting model should discriminate, intentionally or unintentionally, against a protected class. Zest's technology ensures that both the data used in a credit underwriting model and its results are consistent and impartial in its decisioning.
Regular monitoring	The ability to continually monitor a model and ensure that it performs consistently over time. Monitoring results may ultimately support a decision to adjust or "refit" a model, which restarts the training and testing cycle.