# A New Machine Learning Approach Answers What-If Questions

**Article** *in* MIT Sloan Management Review · February 2025



Yash Raj Shrestha
University of Lausanne

**A New Machine Learning Approach Answers What-If Questions**

**Cite as:** Feuerriegel, S., Shrestha, Y. R. & von Krogh, G. (2025). **A new machine learning approach answers "what if" questions.** MIT Sloan Management Review

Causal ML enables managers to explore different scenarios to improve decision-making.

By Stefan Feuerriegel, Yash Raj Shrestha, and Georg von Krogh

Machine learning is now widely used to guide decisions in processes where gauging the probability of a specific outcome — such as whether a customer will repay a loan — is sufficient. But because the technology, as traditionally applied, relies on correlations to make predictions, it offers managers no insight when it comes to understanding the impact of different choices on business outcomes.[1]

Consider an R&D manager at a large company who is faced with deciding how much to invest in a new technology. Using traditional ML, they can ask what will happen when they increase their spending. They might find a strong correlation between higher levels of investment and higher revenue when the economy is growing. And they might conclude that, since economic conditions are favorable, they should increase the R&D budget.

But should they, really? If so, by how much? External factors, such as levels of consumer spending, technology spillover from competitors, and interest rates, also influence revenue growth. Comparing how different levels of investment might affect revenue while considering these other variables is useful for the manager to determine an R&D budget that delivers the greatest benefit to the company.

Causal ML — an emerging area of machine learning — can help to answer such what-if questions through causal inference. Similar to how marketers use A/B tests to infer which of two ads is likely to generate more sales, causal ML can inform what might happen if managers take a particular action. [2]

This makes the technology useful in many of the same business functions that use traditional ML, including product development, manufacturing, finance, human resources, and marketing.[3] Traditional ML is still the go-to approach when making predictions--such as forecasting stock prices or recommending products that customers are most likely to buy—is the only goal. When companies want to predict what would happen if they made one decision or another—such as whether a 10% discount or none is more likely to induce a customer to make a repeat purchase, they need causal ML.

Our research on AI and our experience helping companies apply causal ML suggests a path to using the technology successfully. (See "The Research.") Companies will need the right expertise, too, and should boost employees' literacy in causal ML.

**What causal ML can and cannot do**

Causal ML is a powerful tool, but managers may find the name misleading. It would be better to call it "counterfactual prediction" to more accurately reflect what it does: predict outcomes based on hypothetical actions. The technology is best understood as a way to make better guesses rather than as a source of definitive answers. Framing it in this way can remind managers not to overinterpret the results.

It does this using causal inference, a method that looks at past results to understand and explain cause-and-effect relationships among variables. But instead of focusing on why something happened, causal ML applies these relationships to predict the effects of interventions in new, forward-looking settings.

However, the method cannot explain why a causal relationship exists between a particular factor and the outcome it affects. For instance, a causal ML model might predict that reducing an R&D budget will decrease revenue, but it will not explain why that relationship exists or whether confounders—factors affecting both the decision and the outcome—might change and invalidate that prediction. Managers should use their domain expertise to evaluate whether a given prediction makes sense. This approach helps ensure that the model's predictions are interpreted correctly and remain relevant to real-world decisions.

Like traditional ML, causal ML is most effective when managers have large volumes of data, their options are clearly defined, and the desired outcome is well-understood. It is generally unsuitable for one-off decisions and in scenarios requiring intuition or creativity.

**Choose the right problem—and data**

Causal ML is best at predicting the outcomes of straightforward decisions that are supported by ample historical data from internal and external sources. Questions about operations can be good candidates because they are made frequently, and companies have a lot of data to support them. .[4] Some examples include:

- Booking.com collects data from thousands of hotel reservations every hour. Marketers at the company use causal ML to determine not only whether to give discounts but also which customers should get them.
- Chocolate maker Lindt has extensive data about environmental conditions, equipment, packaging, and other factors that affect the quality of its world-famous truffles. Manufacturing managers use causal ML to help them fine-tune parameters such as the temperature of machines and the configurations of truffle molds..[5]
- ABB Hitachi turned to causal ML to reduce failure rates in its semiconductor manufacturing process. Using machine performance data, the company was able to cut its yield loss by about half by identifying which combination of machines consistently produced the best quality chips.[6]

At Novartis, managers who were trained to distinguish among different kinds of machine learning were able to identify several decision-making tasks where replacing traditional ML with causal ML offered significant benefits. The company had used traditional ML to forecast sales, but these predictions were not helping them with how to allocate their marketing budget. With their knowledge of causal ML, managers decided to evaluate how different promotional campaigns would affect future sales. They used the predictions to distribute resources to the campaigns that were likely to be most effective.

A decision that is suitable for causal ML can be expressed as a number or a binary choice (such as an amount of revenue or buy/hold). It may also be framed as a question about which action to take: to allocate a marketing budget of $10,000 or $15,000 for the next quarter, to offer a 10% discount or none on a product.[7]

Further, causal ML cannot effectively address every potential use case, even if it is suitable on the surface. Confounders introduce biases that affect predictions if they are not accounted for. These can be challenging or impossible to test and affect the accuracy of predictions. If, for example, data is only available for product sales during an economic upturn, predictions of product sales during a downturn would be less reliable.

When managers have determined what they want to decide, how they will measure the outcome, and affirmed they have enough data, they can begin to work with data scientists to assemble and categorize that data to build their causal ML model. Business leaders and other individuals with domain knowledge are essential partners to data scientists and machine learning experts in building causal ML models that provide reliable results.

Training the model to capture complex cause-and-effect relationships requires data from at least a few dozen—and ideally, hundreds or thousands—of historical decisions. With massive amounts of data, the model can uncover connections between variables that may be unknown to managers or difficult to quantify. Less data leads to less accurate predictions.

Broadly, causal ML requires three categories of data alluded to above: decisions, outcomes, and confounders. Decision data encompasses what managers have done in the past, such as the staffing or budgets they set, discounts they offered, investments they made, or processes they changed. Outcome data may include any measurable business result, such as sales volume, revenue growth, quality metrics, or productivity.

Confounders – the variables that affect both the outcome and the decision – come from internal or external sources. They may include economic conditions, workforce composition, and competitor behavior, and they can vary with the decision being made. For a marketing decision, the type of device customers use may be a confounder because those with more expensive smartphones may tend to spend more money whether or not they respond to an incentive.

For example, Neue Zürcher Zeitung, an international media company that publishes the largest circulation newspaper in Switzerland, implemented causal ML to improve the effectiveness of editors' content promotion decisions. The decision variable was whether an online article was promoted on one of two front pages they served to readers. The outcome variable was a performance score that combined website traffic, reader engagement, and subscription signups. Confounders included time factors (such as the hour of the day), content characteristics (such as the article format), past performance indicators (including clicks), and past promotion decisions (including whether the article had been promoted elsewhere).

**Identify Possible Causal Factors**
A valuable lesson from our work has been to sketch a causal graph on a whiteboard that illustrates the expected relationships between the outcome, the decision, and the confounders at the start of the model development process. Managers' knowledge and expertise are essential here because they have repeatedly made decisions and learned to anticipate certain results.

The causal graph tells the data scientists (who should be experts in causal inference) whether to treat a variable as a "cause" or an "effect" in the model. In this way, the team can rule out reverse causality errors. That is, they will ensure the model does not misinterpret one variable as causing another when, in reality, the effect is the opposite.

Imagine a celebrity with millions of social media followers. If we don't know much about social media or stardom, we might conclude that fame comes from having a high follower count. The opposite is more likely to be true. As even the average teenager has observed, a person has to do something to get noticed before millions of strangers will follow their TikTok account.

Considering the example of R&D spending, the budget influences revenue, never the other way around. Meanwhile, confounders such as the economic climate, market trends, or team expertise are acknowledged as driving both the budget decision and business outcomes but are not influenced by either. The model will take all this into account. (See, "A causal graph for an R&D budget decision.")

**Choose the output**

Next, managers need to choose the type of answer the model will give for the question. (in statistics, the output, or estimand). They can predict the end result of a decision or on the relative benefit of one alternative compared to another.

Each can be useful depending on how the manager is thinking through a decision. Focusing on end results helps with strategic planning, such as predicting potential revenues under different budget scenarios or personalizing incentives for individual customers. However, comparing the incremental effects of different decisions is often sufficient for making one: If a manager wants to know which of two ads is likely to boost sales more effectively, they do not necessarily need to predict how much revenue each variant might generate. They only need to know the relative benefit: that one ad is likely to generate three times more revenue than the other. Moreover, focusing on relative benefits generates more reliable predictions than focusing on end results. We recommend pursuing only as much granularity as necessary.

Neue Zürcher Zeitung editors were interested in predicting the actual click rates for each article they promoted, but the company opted instead to predict the net gain in performance from promoting an item. This approach enabled causal ML to make more accurate predictions about which content, when promoted, would increase clicks and subscriptions. Editors learned that promoting articles by the editor-in-chief significantly increased both outcomes.[8] They had promoted the top editor's articles sparingly, and the findings served as a starting point to revise their promotional strategy.

**Train, test, and validate the model**

Once managers have defined the decision they are making and their preferred output, data and machine learning scientists can choose a causal ML model that is right for the job. When they have implemented it, machine learning engineers will train the model using the previously categorized data.

The final step is to test and validate the causal ML model in practice to ensure it is reliable and that its predictions improve business performance. Validation also offers the opportunity for decision-makers, including senior leaders, to gain trust in its predictions. Starting with relatively simple and straightforward problems where clear decision alternatives can be identified and assessed makes this step easier to accomplish.

Testing and validation require care because managers can only observe the outcome of the decision that was made in the real world. They have no way to know the outcome had a different decision been made. Two strategies, “human in the loop” and the familiar A/B testing, have proven successful.

Neue Zürcher Zeitung chose to integrate the model's recommendations into human decision-making processes.[9] Their causal ML model recommended which content to promote while the editors made the final decisions. The model relied on the same information that editors used in their decisions previously; they could trust that causal ML was not missing some element. Typically, causal ML made recommendations that matched editors’ own gut feelings, which gave them confidence that the model was reliable.

Some decisions are tricky, and editors know their judgment isn’t perfect. In cases where causal ML recommends a different decision than they would have made, the editors can test the recommendation and see the result. Over time, they can see that causal ML makes reliable recommendations in ambiguous situations. They can follow the causal ML recommendations instead of their instincts more frequently.

ABB Hitachi used A/B testing to validate the causal ML models they built to improve manufacturing quality. In one application, managers used the model to predict which of several machines would produce the best quality output in the etching and implantation steps of the semiconductor fabrication process and contribute to the highest quality overall.

To confirm the predictions were reliable, managers did a controlled experiment in which they changed the machine used for etching and implantation and kept the machines used for other processes the same. They found that the better machine for etching and implantation was the same one that the causal ML model had predicted. Thanks to causal ML, managers found and addressed the source of manufacturing issues more efficiently than with either manual methods or traditional ML.[10]

**Prepare the organization**

While causal ML has the potential to improve decisions, implementing such systems requires a high level of AI literacy in the workforce, specialized technical expertise, and patience—because these projects may take longer to develop than traditional ML applications. Managers can prepare their organizations by educating themselves and their workforce about causal AI and building the interdisciplinary teams needed to develop their applications.

Many companies today are investing heavily in educating employees about traditional ML and generative AI (such as ChatGPT) to stay competitive and innovative. If the organization plans to use causal ML, it needs to include this technology in its AI literacy efforts. Employees who are alert to the strengths and limitations of different AI approaches will be empowered to find opportunities to use them effectively.

We found that to excel at using causal ML, teams need strong expertise in data science and machine learning, along with domain knowledge. However, building such teams can be costly, particularly when companies need to hire data scientists or turn to external consultants and partners.

Moreover, data scientists and machine learning engineers are typically assigned to different teams. They need to work closely when developing and implementing causal ML models, with strong engagement from business stakeholders who have domain knowledge (Domain knowledge is also essential in traditional ML but is often less rigorously applied because teams do not deeply consider the underlying relationships between variables when building those models.)

For example, at Neue Zürcher Zeitung, editors and marketers have insight into editorial processes, customer preferences, and long-term objectives of the brand that help data scientists define variables that measure these. At ABB Hitachi, engineers supplied the insight to define which production variables to include in the models.

Interdisciplinary teams are often plagued by a lack of common understanding, vocabulary, and ways of working. Managers need to foster an environment where cross-functional collaboration can thrive, and all relevant stakeholders are involved throughout the model development process. Regular workshops, meetings, and training sessions where data scientists, machine learning engineers, and domain experts jointly explore problems, refine models, and discuss the implications of the findings together can foster an environment where cross-functional collaboration thrives.

## Conclusion

Machine learning has changed how numerous organizations make decisions; causal ML can deepen insights further by predicting the effects of different choices on business outcomes. Companies are more likely to gain from machine learning when decision-makers trust the results. Knowing what causal ML can do and how it compares to traditional ML can help them choose the right projects for both technologies and increase the success rate.

When managers use causal ML prudently to explore the options for straightforward decisions, they can significantly improve their operations—and, ultimately, their financial results.

### The Research

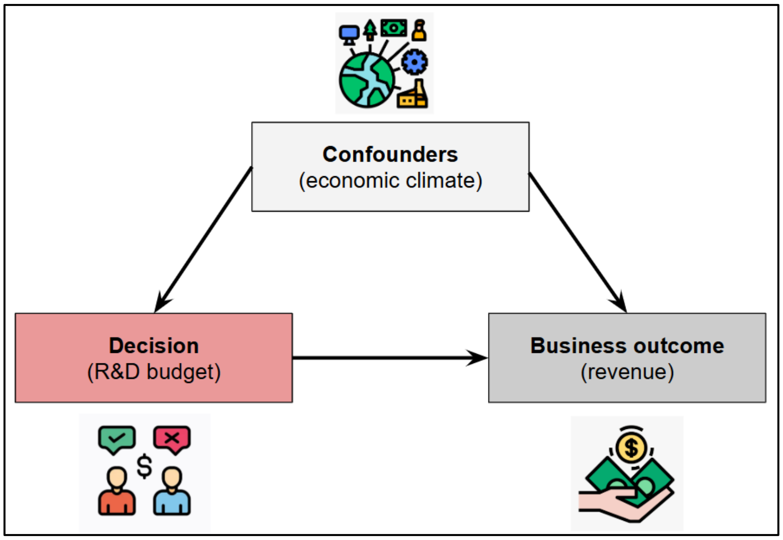
- The authors worked with companies in manufacturing, pharmaceuticals, finance, travel, and media to develop and implement causal ML models for a variety of business functions. The companies included ABB Hitachi, Booking.com, Ethon AI, Neue Zürcher Zeitung, Novartis, and UBS.

- Over five years, they implemented and evaluated the models and documented the practices that contributed to the companies' success using the technology.

## [FIGURE]A causal graph for an R&D budget decision

[caption] A causal graph helps data scientists understand whether a variable should be treated as a "cause" or an "effect" in a causal ML model. The graph describes the role of different variables and how they are expected to interact. The arrows indicate the direction of influence for each. Here, the decision is the R&D budget allocation. It

influences the business outcome (revenue). The economic climate is a confounder that influences both the decision and the outcome.




## Endnotes

[1] S. Feuerriegel, Y. Shrethsa, G. von Krogh, et. al., "Bringing Artificial Intelligence to Business Management," Nature Machine Intelligence 4 (July 2022): 611-613. P. Hünermund, J. Kaminski, and C. Schmitt, "Causal Machine Learning and Business Decision Making," SSRN (February 19, 2022).

[2] S. Feuerriegel, D. Frauen, V. Melnychuk, et. al., "Causal Machine Learning for Predicting Treatment Outcomes," Nature Medicine 30 (April 2024): 958–968; V. Chernozhukov, C. Hansen, N. Kallus, et. al., "Applied Causal Inference Powered by ML and AI" (CausalML-book.org: July 28, 2024); and C. Fernández-Loría and F. Provost, "Causal Decision Making and Causal Effect Estimation Are Not the Same … and Why It Matters," INFORMS Journal on Data Science 1, no. 1 (April-June 2022): 4-16.

[3] M. von Zahn, K. Bauer, C. Mihale-Wilson, et. al., "Smart Green Nudging: Reducing Product Returns Through Digital Footprints and Causal Machine Learning," Marketing Science (August 8, 2024); E. Ascarza, "Retention Futility: Targeting High-Risk Customers Might Be Ineffective," Journal of Marketing Research 55, no. 1 (February 2018): 80-98; J. Yang, D. Eckles, P. Dhillon, et. al., "Targeting for Long-Term Outcomes," Management Science 70, no 6 (June 2024): 3841-3855; and M. Kraus, S. Feuerriegel, and M. Saar-Tsechansky, "Data-Driven Allocation of Preventive Care with Application to Diabetes Mellitus Type II," Manufacturing & Service Operations Management 26, no.1 (January-February 2024): 137-153.

[4] G. von Krogh, S. Ben-Menahem, and Y. Shrestha, "Artificial Intelligence in Strategizing: Prospects and Challenges," in "Strategic Management: State of the Field and Its Future," eds. I. Duhaime, M. Hitt and M. Lyles. (New York: Oxford University Press, 2021).

[5] ETH AI Center, "Premium Chocolate Production Perfected: AI's Role in Quality Excellence," 2023.

[6] J. Senoner, T. Netland, and S. Feuerriegel, "Using Explainable Artificial Intelligence to Improve Process Quality: Evidence from Semiconductor Manufacturing," Management Science 68, no. 8 (August 2022): 5704-5723.

[7] H. Wasserbacher and M. Spindler, "Machine Learning for Financial Forecasting, Planning and Analysis: Recent Developments and Pitfalls," Digital Finance 4 (March 2022): 63-88.

[8] J. Persson, S. Feuerriegel, and C. Kadar, "Off-Policy Learning for Audience-Wide Content Promotions," working paper, 2023.

[9] Persson, 2023.

[10] Senoner, 2022.