INSIGHTS · WHAT AN INSIGHT IS

Predictive Insights versus Explanatory Insights: Different Knowledge Purposes

Some insights tell an organisation what will happen; others tell it why. This article explains the difference between predictive and explanatory insight, why the best model for one is often not the best for the other, and how to choose and combine them for business decisions.

ModQu research article · 1,958 words · 20 references (APA 7th edition)

Two questions sit behind most requests for insight. The first is "What will happen?" Which customers will leave next quarter, how much will a new product sell, which applicants will repay a loan? The second is "Why does it happen?" Why do customers leave, what drives demand, what makes some borrowers default? The answers to these questions are both called insights, but they are different kinds of knowledge, produced by different methods and judged by different standards. Confusing them is one of the most common sources of disappointment in analytics.

Two purposes, two standards

Shmueli (2010) set out the distinction clearly in an influential paper titled "To explain or to predict?" *Explanatory* modelling aims to test or estimate causal relationships: to find out whether, and how much, a factor such as price or service quality affects an outcome such as loyalty. *Predictive* modelling aims to forecast new or future observations as accurately as possible, regardless of whether the model reflects the true causes. Shmueli showed that the two goals lead to different choices at every stage of analysis, from the variables collected to the way models are evaluated.

The standards of success differ too. An explanatory insight is judged by whether its causal claim is correct: whether the factor really has the effect claimed, in the direction and size claimed. A predictive insight is judged by its accuracy on data it has not seen. A model can predict well without explaining anything, and it can explain correctly while predicting poorly.

Philosophers once argued that explanation and prediction were two sides of the same coin. Hempel and Oppenheim (1948) proposed that a good scientific explanation has the same logical form as a prediction: the conditions and laws that explain an event would have allowed it to be predicted in advance. In practice, the two have drifted apart, because the methods that produce the most accurate predictions are often not the ones that reveal causes.

The two cultures

Breiman (2001) described two cultures in statistics. The *data modelling* culture starts by assuming a model of how the data were generated, usually a relatively simple equation, and then estimates its parameters. Its aim is to understand the process. The *algorithmic modelling* culture treats the process as unknown and searches for whatever function predicts the outcome best, using methods such as decision trees, random forests and neural networks. Breiman argued that statisticians had relied too heavily on the first culture and had ignored the predictive power of the second.

Machine learning has since made the algorithmic culture dominant in many applied fields. Mullainathan and Spiess (2017) explained the difference for economists: machine learning excels at predicting an outcome, but the coefficients of a machine-learning model should not be interpreted as estimates of causal effects. Yarkoni and Westfall (2017) argued that psychology, long focused on explanation, would benefit from paying more attention to prediction, since many explanatory models fail to predict new data well, a sign that they may be overfitted to the samples on which they were developed.

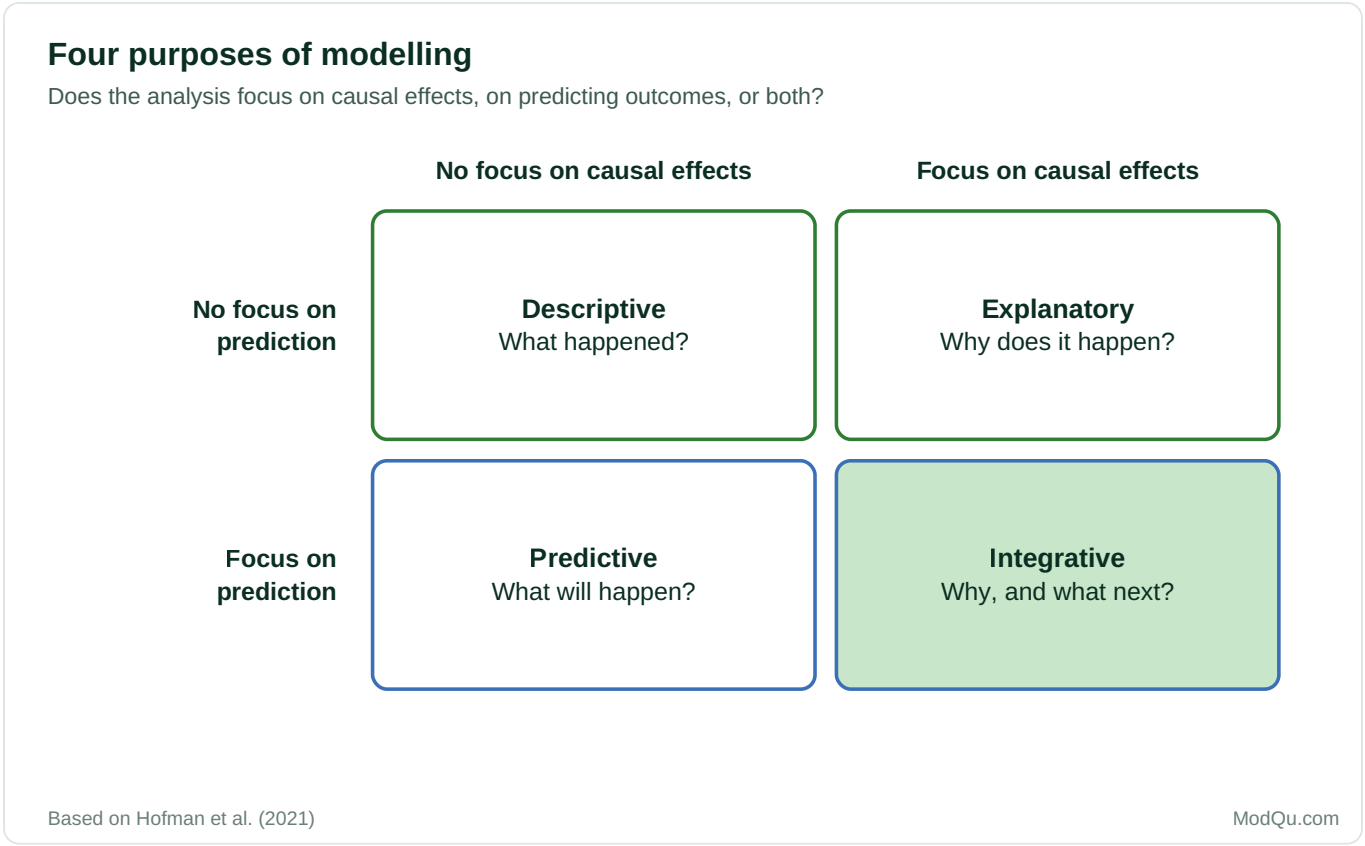


Figure 1. Four kinds of modelling purpose. Most business insight is either explanatory (why) or predictive (what next); the most demanding work integrates both. Diagram by ModQu, based on Hofman et al. (2021).

Why the best predictor is not always the best explanation

Several features of data and models explain why the two goals pull apart.

Correlated proxies. A predictive model will use any variable that improves accuracy, including variables that are only proxies for the true causes. A model that predicts customer churn from the number of help-desk calls may predict well, but the calls are a symptom of dissatisfaction, not its cause. Reducing the number of calls, for example by making the help desk harder to reach, would not reduce churn.

Complexity and interpretability. Accurate predictive models often combine hundreds of variables in complex ways that resist simple explanation. Explanatory models are usually simpler, because their purpose is to isolate the effect of a few factors.

Stability over time. A causal relationship should hold when circumstances change, as long as the causal mechanism remains. A predictive relationship may hold only as long as the patterns in the training data continue. Lazer, Kennedy, King and Vespignani (2014) documented how Google Flu Trends, which had predicted influenza levels from search queries (Ginsberg et al., 2009), later overestimated flu prevalence substantially. Changes in search behaviour and in the search engine itself broke the patterns the model relied on. The model had predicted without explaining, and when the circumstances changed it had no way of adjusting.

Simplicity can be enough. In many predictive tasks, simple models do almost as well as complex ones. Dawes (1979) showed that even "improper" linear models, which give equal weight to each relevant variable, often predicted outcomes better than expert judgement. Hand (2006) argued that much of the claimed improvement from sophisticated classification methods disappears in real applications, where data change and measurement is imperfect.

When prediction is what matters

Some decisions need only a good prediction. Kleinberg, Ludwig, Mullainathan and Obermeyer (2015) called these *prediction policy problems*. Their example contrasts two decisions. Someone deciding whether to carry an umbrella needs to know whether it will rain, not why. Someone deciding whether to perform a rain dance needs to know whether the dance causes rain. Many business decisions are of the umbrella type: how much stock to order, which customers to contact first, which transactions to flag for review. For these, predictive insight is sufficient, and the more accurate the better.

Evidence from decades of research supports relying on systematic prediction in such cases. Grove, Zald, Lebow, Snitz and Nelson (2000) combined 136 studies comparing clinical judgement with mechanical, formula-based prediction in fields such as medicine and psychology. Mechanical prediction was as accurate as, or more accurate than, expert judgement in the large majority of comparisons. In marketing, Morwitz, Steckel and Gupta (2007) found that stated purchase intentions predict sales, but with a systematic bias that can be corrected once it is measured, which is a predictive rather than an explanatory use of survey data.

When explanation is what matters

Other decisions require knowing what will happen *if the organisation acts*. Will a price cut increase revenue? Will a new onboarding process reduce churn? Will a change in a product's design improve satisfaction? These are rain-dance decisions. A model that merely predicts which customers will churn cannot say which intervention will change their

behaviour. Answering that requires explanatory insight, ideally from experiments or from careful causal analysis (Pearl, 2009). Athey (2017) emphasised that many policy questions need estimates of causal effects, not just predictions, and that big data alone do not supply them.

Using predictive models for explanatory purposes can cause harm. Obermeyer, Powers, Vogeli and Mullainathan (2019) studied an algorithm widely used in United States hospitals to identify patients who needed extra care. The algorithm predicted future health-care costs accurately. But costs were used as a stand-in for health needs, and because less money had historically been spent on Black patients with the same level of illness, the algorithm systematically underestimated their needs. The prediction was accurate; the implied explanation, that cost reflected need, was wrong.

A worked example: customer churn

A subscription business asks its analytics team for "insight into churn". The team builds two kinds of analysis.

The first is a predictive model. Using account records, it estimates each customer's probability of cancelling in the next 90 days. The strongest predictors are a recent fall in usage, a recent complaint, a recent payment failure and the absence of any login in the past fortnight. Tested on a holdout sample of customers whose outcomes were already known, the model ranks churners well: the 10% of customers with the highest scores account for a large share of actual cancellations. This is a valuable predictive insight. It tells the retention team whom to contact first.

The second is an explanatory analysis. The team wants to know what would actually keep customers. It runs an experiment in which randomly selected high-risk customers receive either a discount, a personal call offering help with the product, or nothing. The call reduces cancellations noticeably; the discount barely changes them, and it reduces revenue from customers who would have stayed anyway. This is an explanatory insight. It tells the business which action works and, by implication, that the main driver of churn among these customers is difficulty using the product rather than price.

Neither analysis could have replaced the other. The predictive model could not say whether a discount or a call would work, because neither had been used systematically in the past. The experiment could not say whom to target, because the business could not afford to call every customer. Varian (2014) made the general point that machine learning and causal methods are complements: one finds patterns that predict, the other tests which levers move outcomes. Used together, they turned a request for "churn insight" into a targeted programme with a measurable effect.

Making predictions explainable

Because organisations often want both accuracy and understanding, researchers have developed methods to explain the predictions of complex models. Ribeiro, Singh and Guestrin (2016) proposed a technique that explains individual predictions by approximating the complex model locally with a simpler one. Such methods show which variables drove a particular prediction. They do not, however, turn predictive associations into causal effects. Rudin (2019) argued that for high-stakes decisions it is better to use models that are interpretable from the start than to explain black-box models after the fact, because post-hoc explanations can be misleading.

Hofman and colleagues (2021) proposed that social science should aim for *integrative* modelling, which combines a focus on specific causal effects with a focus on predicting outcomes accurately. Explanatory models should be tested on how well they predict new data, and predictive models should be examined for what they reveal about mechanisms. Box (1976) made a related point decades earlier: all models are simplifications, and their value lies in being useful for a specific purpose.

A practical guide for insight teams

1. **Ask what decision the insight will inform.** If the decision depends only on what will happen, a predictive insight is enough. If it depends on what will happen *if* the organisation intervenes, an explanatory insight is needed.
2. **Label insights by type.** A slide that shows "the top drivers of churn" from a predictive model invites a causal reading. Say clearly whether a relationship is predictive or causal.
3. **Judge each by its own standard.** Evaluate predictive insights on accuracy with new data, using holdout samples or later outcomes. Evaluate explanatory insights by the strength of the causal design: experiments, natural experiments or careful control of alternative explanations.
4. **Watch for proxies.** When a predictor seems actionable, ask whether changing it would change the outcome, or whether it is only a symptom.
5. **Monitor predictive models over time.** Patterns change. A model that predicted well last year may fail this year, as Google Flu Trends did. Regular checks against actual outcomes are essential.
6. **Combine the two where possible.** Use predictive models to identify who is likely to act, and experiments to discover what would change their behaviour.

Key takeaways

Predictive insights tell an organisation what is likely to happen; explanatory insights tell it why, and what would happen if it acted. The two serve different decisions and are judged by different standards. The most accurate predictive models often rely on proxies and complex combinations of variables that do not reveal causes, while sound explanatory models may predict less accurately. Treating a prediction as an explanation can lead to ineffective or harmful actions. Insight teams should match the type of insight to the decision, label it clearly and, where the stakes justify it, combine accurate prediction with credible causal evidence.

References

Athey, S. (2017). Beyond prediction: Using big data for policy problems. *Science, 355*(6324), 483–485. https://doi.org/10.1126/science.aal4321

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association, 71*(356), 791–799. https://doi.org/10.1080/01621459.1976.10480949

Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science, 16*(3), 199–231. https://doi.org/10.1214/ss/1009213726

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34*(7), 571–582. https://doi.org/10.1037/0003-066X.34.7.571

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature, 457*(7232), 1012–1014. https://doi.org/10.1038/nature07634

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12*(1), 19–30. https://doi.org/10.1037/1040-3590.12.1.19

Hand, D. J. (2006). Classifier technology and the illusion of progress. *Statistical Science, 21*(1), 1–14. https://doi.org/10.1214/088342306000000060

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science, 15*(2), 135–175. https://doi.org/10.1086/286983

Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M. J., Vazire, S., Vespignani, A., & Yarkoni, T. (2021). Integrating explanation and prediction in computational social science. *Nature, 595*(7866), 181–188. https://doi.org/10.1038/s41586-021-03659-0

Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review, 105*(5), 491–495. https://doi.org/10.1257/aer.p20151023

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science, 343*(6176), 1203–1205. https://doi.org/10.1126/science.1248506

Morwitz, V. G., Steckel, J. H., & Gupta, A. (2007). When do purchase intentions predict sales? *International Journal of Forecasting, 23*(3), 347–364. https://doi.org/10.1016/j.ijforecast.2007.05.015

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives, 31*(2), 87–106. https://doi.org/10.1257/jep.31.2.87

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511803161

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144). ACM. https://doi.org/10.1145/2939672.2939778

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, *25*(3), 289–310. https://doi.org/10.1214/10-STS330

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28. https://doi.org/10.1257/jep.28.2.3

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science, 12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393