

MACHINE LEARNING AND FINTECH

Spencer McManus*

CITE AS: 1 GEO. L. TECH. REV. 428 (2017)

<https://perma.cc/AN4A-H6YG>

INTRODUCTION.....	428
A BRIEF HISTORY OF MACHINE LEARNING	428
PATTERN RECOGNITION AND EXPLANATION-BASED LEARNING: AN E-COMMERCE EXAMPLE.....	430
Training Data	430
Decision Trees	431
Evaluating the Model	432
EBL AND FRAUD DETECTION.....	437
CONCLUSION.....	435

INTRODUCTION

Machine learning describes the process through which computers can learn without continued human input. In the era of big data, machine learning is particularly promising because it allows for identification of patterns in large data sets. Machine learning has applications in fields as diverse as medicine, e-commerce, and banking. This essay will discuss the application of machine learning, particularly explanation-based learning, to the financial tech industry, focusing on fraud detection.

A BRIEF HISTORY OF MACHINE LEARNING

The concept of machine learning first arose in 1950 with Alan Turing’s paper *Computing Machinery and Intelligence*,¹ in which Turing proposed to answer the question, “Can machines think?” To answer this question, Turing crafted what became known as the “Turing Test” with three participants: one human judge, one human player, and one computer. The judge, placed separately from the human and the computer, aims to determine which of the two is a human and which is a computer. A computer “passes” the Turing Test when the judge cannot consistently distinguish the computer

* GLTR Staff Member; Georgetown Law, J.D. expected 2018; University of California, Davis, B.A.S. 2015. © 2017, Spencer McManus.

¹ A.M. Turing, *Computing Machinery & Intelligence*, 59 MIND 433 (1950).

player from the human player.² Turing predicted that humans could program computers that would pass the test by 2000.³

Over the next four decades, scholars and programmers refined the concept of machine learning and developed new tests. In 1959, IBM programmer Arthur Samuel created a checkers program in which the computer improved progressively the more it played.⁴ Programmers focused on developing machines that performed pattern recognition over the next two decades. These efforts culminated in the introduction of Explanation-Based Learning (“EBL”),⁵ in which a machine uses a set of programmer-supplied “training data” to identify patterns, synthesize rules, and apply the rules to new sets of data.⁶

From the 1990’s to today, work has transitioned to developing machines that can handle large amounts of data to draw conclusions.⁷ Machine learning has been extended to include “deep learning,” which involves use of increased processing power to analyze visual and auditory data in real-time.⁸ Large technology companies have developed their own proprietary machine learning code that acts as the backbone for certain features of their products.⁹ Future development focuses on continued improvement in natural language processing—which allows for human voice interaction with devices¹⁰—and applying machine learning to new industries.

² *Id.* at 442.

³ *Id.*

⁴ See Bernard Marr, *A Short History of Machine Learning Every Manager Should Read*, FORBES (Feb. 19, 2016, 2:31 AM), <https://www.forbes.com/sites/bernardmarr/2016/02/19/a-short-history-of-machine-learning-every-manager-should-read/#67840fd815e7> [<https://perma.cc/CGD3-DEYE>].

⁵ See *infra* Part II.

⁶ See generally Gerald Dejong & Raymond J. Mooney, *Explanation-Based Learning: An Alternative View*, 1 MACH. LEARNING 145 (1986).

⁷ See Marr, *supra* note 4.

⁸ See Robert D. Hof, *Deep Learning*, MIT TECH. REV., <https://www.technologyreview.com/s/513696/deep-learning/> [<https://perma.cc/9F68-THTJ>].

⁹ Examples include Facebook’s DeepFace, which powers the social network’s facial detection feature, and numerous digital assistant applications, including Apple’s Siri, Amazon’s Alexa, and Microsoft’s Cortana. See, e.g., Steven Levy, *The iBrain is Here, and It’s Already Inside Your Phone*, BACKCHANNEL (Aug. 24, 2016), <https://backchannel.com/an-exclusive-look-at-how-ai-and-machine-learning-work-at-apple-8dbfb131932b#.6wi4d8qcy> [<https://perma.cc/Z9KG-8LQ8>] (explaining how Apple uses machine learning in their products, including Siri).

¹⁰ See Perry Li, *Natural Language Processing*, 1 GEO. L. TECH. REV. 98 (2016), <https://www.georgetownlawtechreview.org/natural-language-processing/GLTR-11-2016/>

PATTERN RECOGNITION AND EXPLANATION-BASED LEARNING: AN E-COMMERCE EXAMPLE

Although there are several methodologies for machine learning, this article focuses on explanation-based learning. Explanation-based learning (“EBL”) involves teaching a machine to detect patterns in data based on a set of programmer-supplied “training data,” using the patterns to create a rule and then applying the rule to larger sets of data to make predictions. A simplified but powerful example from the e-commerce industry will help illustrate the process.¹¹

Retail companies face the challenge of catering to individual customers in a growing global economy. Machine learning can help retailers by providing extremely personalized predictions about how an individual’s shopping habits may change given a change in personal circumstances. The simple system illustrated here will involve a machine learning system predicting whether a customer is pregnant.

Training Data

The first requirement for a machine learning system is “training data.” Training data consists of different data points (called “features”), which come together to form an individual “record,” and an output value (the “target”).¹² Training data is necessary because the machine cannot make predictions without examples of how the different features affect the output. In our example, the features will be the customer’s age and whether or not she purchases two products commonly associated with pregnancy. These features come together to form ten “records”: in this case, a purchasing history for one

[<https://perma.cc/S2AU-SKZL>]; see generally Winfred Phillips, *Introduction to Natural Language Processing*, CONSORTIUM ON COGNITIVE SCI., INSTRUCTION (2006), http://www.mind.ilstu.edu/curriculum/protothinker/natural_language_processing.php [<https://perma.cc/VLB9-JNHJ>].

¹¹ This example is based on the (in)famous Target “baby club” story, in which Target, using a machine learning model, predicted that a teenage customer was pregnant. Target started sending her baby coupons, which were discovered by her father, who had not yet been informed about the pregnancy. For background, including more on how e-commerce companies leverage machine learning, see Charles Duhigg, *How Companies Learn Your Secrets*, N.Y. TIMES MAGAZINE (Feb. 16, 2002), <https://nyti.ms/2jEboTD> [<https://perma.cc/9Y2L-HCLH>].

¹² Michael Manapat, *A Primer on Machine Learning for Fraud Detection*, STRIPE, <https://stripe.com/radar/guide> (last visited Feb. 5, 2017) [<https://perma.cc/D2QT-FWPS>].

customer. The target is whether the customer was actually pregnant.¹³ Table 1 shows this data.

#	AGE	PREGNANCY TEST?	PRE-NATAL SUPPLEMENTS?	<i>PREGNANT?</i>
1	24	Yes	Yes	Yes
2	36	Yes	No	No
3	18	No	No	Yes
4	40	No	Yes	Yes
5	25	No	Yes	Yes
6	33	Yes	No	Yes
7	25	Yes	Yes	Yes
8	32	No	No	No
9	35	No	No	No
10	20	Yes	Yes	No

TABLE 1. Training data for hypothetical pregnancy prediction.

Decision Trees

From this set of ten records and their corresponding outputs, the computer can form a “decision tree,” a process for evaluating the probability of the output occurring given the value of each feature. This is a decision tree for this problem, with row (A) showing the probabilities of pregnancy given the training data in Table 1.

¹³ This hypothetical example ignores the difficulty (and potentially legality) of determining if a customer was actually pregnant. *See id.*

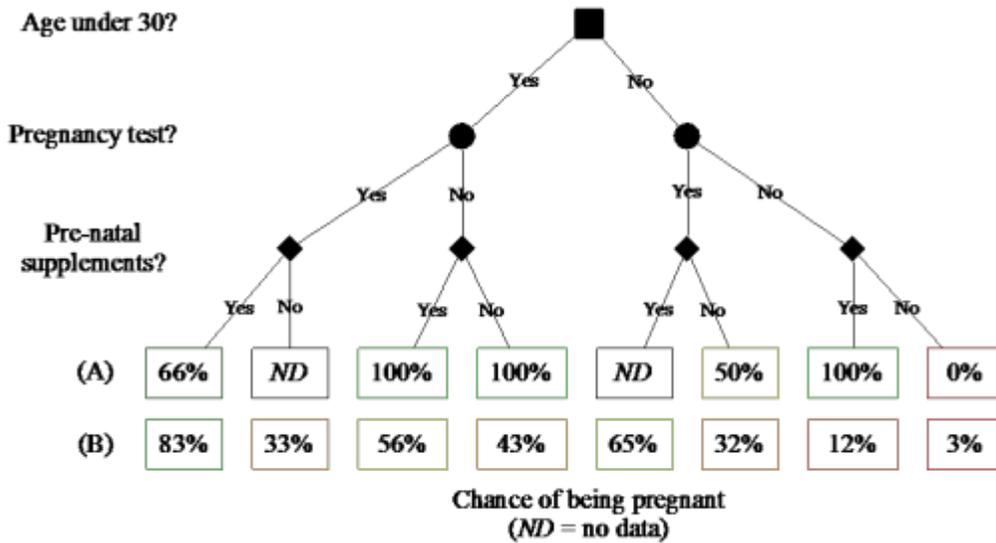


FIGURE 1. Decision tree for data in TABLE 1. © Spencer McManus, 2017.

With this small amount of data, the decision tree is not particularly useful to the retailer. The ten records do not capture the purchasing trends of the entire customer base; two branches of the tree remain empty. Imagine instead that a larger set of training data (say, with 10,000 records) produced the probabilities in row (B) in Figure 1. This data would be useful to a retailer, especially where the data produced high or low probabilities.

Evaluating the Model

Now that the model has been developed from an adequate training set, the retailer can utilize the model to make predictions about new customers. A retailer could, say, send coupons for baby products to a potentially pregnant customer who fits in one of the high probability categories. If the customer is indeed pregnant, the coupons might encourage her to shop at the retailer.

However, no decision tree is perfect because of practical limitations in data collection. In this case, the decision tree uses a limited set of data to produce *probabilities* that a customer with a given shopping history is pregnant. The retailer needs to evaluate if its model is actually effective at predicting if a customer is pregnant.

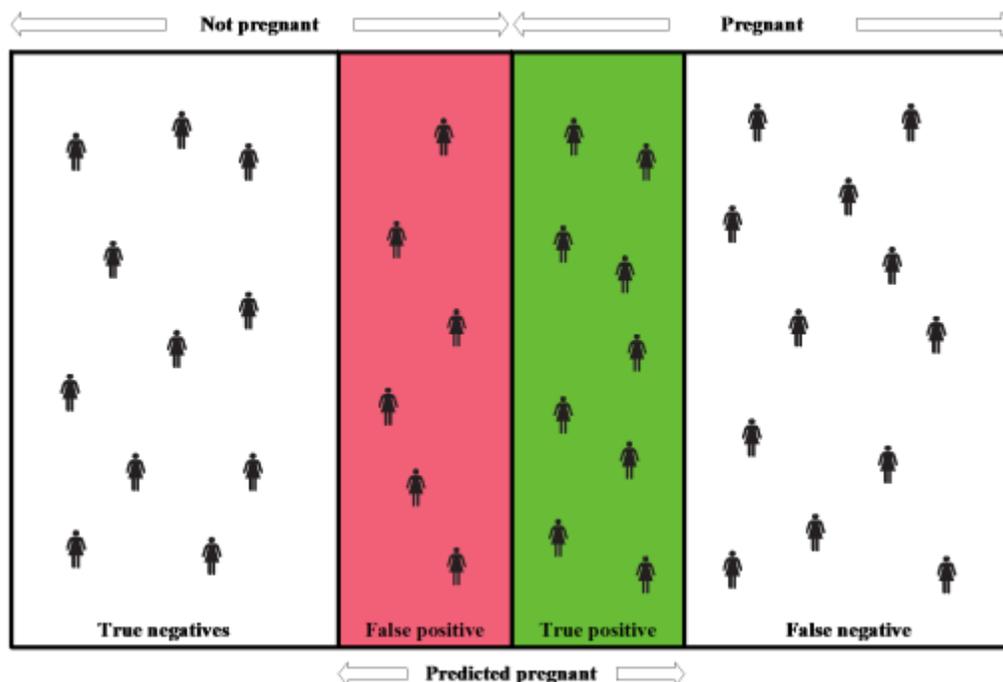


FIGURE 2. Evaluating the efficacy of a machine learning program. © Spencer McManus, 2017.

The shaded areas are where the retailer uses the model and determines that a customer is pregnant; the white areas are where the retailer determined the customer was not pregnant. The retailer should aim to minimize the number of women in the red shaded area (a “false positive,” where the retailer determined someone was pregnant, but she wasn’t) and maximize the number of women in the green shaded area (where the retailer correctly identified someone as pregnant).

The retailer can evaluate this by calculating “precision” and “recall.”¹⁴ In our example, precision is the percentage of customers predicted to be pregnant who actually are. Higher precision indicates fewer false positives. Recall is the percentage of all pregnant customers who are identified by the model. Higher recall indicates fewer false negatives. There is a relationship between precision and recall. As a retailer raises the probability threshold for predicting someone is pregnant, it will reduce false positives (and thus increase precision), but it will also increase false negatives (thus reducing

¹⁴ See Tom Fawcett, *An Introduction to ROC Analysis*, 27 PATTERN RECOGNITION LETTERS 861, 865 (2006).

recall). If a retailer decides that the model must have 85% certainty that a customer is pregnant instead of 75%, it will exclude customers whose product purchases suggest between a 75% and 85% probability of pregnant. It is very possible that customers in this range are pregnant, but that the increased probability threshold will produce false negatives for these customers.

A retailer faces obvious obstacles in determining numbers of actually pregnant customers, but this could be accomplished through surveys of customers. By further refining the model through evaluation of the most predictive features, the trade-off between precision and recall can be reduced, creating a higher quality model and giving the retailer the maximum benefits of a machine learning system.

EBL AND FRAUD DETECTION

EBL is commonly used in the financial technology space to detect credit card fraud. Financial institutions often license fraud-detection software from third-parties. This software, in its most simplified form, utilizes hundreds or thousands of features to form a decision tree, producing probabilities used to predict if a transaction is fraudulent.¹⁵

Using a system similar to that in the pregnancy example, fraud detection companies identify features that, when analyzed together, are highly predictive of fraud. In this simplified example, a fraud detection company could build a system using three different features to detect basic instances of fraud on a single card: the country of use for a charge, the charge amount, and the number of countries used in a given time period.¹⁶

#	COUNTRY	CHARGE AMOUNT	NO. COUNTRIES IN 24-HOUR PERIOD	FRAUD?
1	USA	\$20	1	No
2	RUS	\$150	2	Yes
3	USA	\$200	2	No
4	CAN	\$10	1	Yes
5	CAN	\$15	1	No

TABLE 2. Training data for a simple fraud detection model.

¹⁵ See, e.g., Manapat, *supra* note 12.

¹⁶ See *id.*

Fraud detection can be quite difficult. From a cursory examination of the training data, there does not seem to be a consistent pattern. Although some transactions may be quite obviously fraud (such as record 2, where a charge was made in a country not present in other records), other patterns are not so evident (such as a correlation between number of countries and fraudulent charges). Machine learning becomes particularly useful in the fraud detection industry because it enables companies to quickly analyze complex sets of data. For example, the developer of this model may determine that the third feature is not particularly predictive of fraud. One benefit of EBL in this space is that companies can identify relevant features and exclude irrelevant ones.¹⁷

EBL continues to grow in other financial spheres as well. Banks use EBL to analyze customer traits (including past defaults, job status, and marital status) to approve or reject loans.¹⁸ Other financial institutions use EBL to power “robo-advisors” that advise customers on allocating investments and financial instruments.¹⁹ In the future, EBL could power new security systems for banking (such as facial recognition) or even finance-specific customer service systems.²⁰

CONCLUSION

Since Alan Turing first hypothesized a thinking machine in 1950, machine learning has developed into a powerful tool. In explanation-based learning, one of the many different types of machine learning, a human provides a set of training data, which includes several features and records, from which a machine extrapolates patterns and creates rules. We encounter these systems every day: in e-commerce and fraud detection, machine learning forms a critical backbone. Future development of EBL will focus on applying the technology to new technologies in the era of big data.

¹⁷ *Machine Learning and Fraud Prevention*, RAVELIN (last visited Feb. 9, 2017), <https://www.ravelin.com/resources/machine-learning-and-fraud-prevention> [https://perma.cc/8QMJ-ACCL].

¹⁸ Daniel Faggella, *Machine Learning in Finance – Present and Future Applications*, TEHEMERGENCE (Aug. 15, 2016), <http://tehemergence.com/machine-learning-in-finance-applications/> [https://perma.cc/982M-TS47].

¹⁹ *Id.*

²⁰ *Id.*