

## Editorial

# Learning Algorithms for Ophthalmologists: A Conceptual Primer

Ariana M. Levin, MD<sup>1</sup> Rebekah H. Gensure, MD, PhD<sup>1</sup> Jeff Pettey, MD<sup>1</sup><sup>1</sup>John A. Moran Eye Center, University of Utah, Salt Lake City, Utah

J Acad Ophthalmol 2019;11:e22–e24.

Recent advances, including the first Food and Drug Administration (FDA)-approved artificial intelligence system for detection of diabetic retinopathy, have brought machine learning algorithms to the spotlight in ophthalmology clinical practice.<sup>1</sup> To determine the potential utility and applicability of these systems, it is important for the ophthalmology community to have a conceptual understanding of machine learning principles. Learning algorithms date back at least to the 1950s. In review of published literature, the earliest available studies of machine learning in ophthalmology dated back to 2002 by Sample et al who described patterns in glaucomatous field defects.<sup>2</sup> One of the earliest studies on machine learning for vitreoretinal diseases was an analysis of genetic predictors of proliferative vitreoretinopathy published in 2009.<sup>3</sup> The vast majority of papers on machine learning in ophthalmology have been published in the past 5 years, on a broad range of topics with an increasing emphasis on image analysis. Machine learning has the potential to facilitate diagnosis and optimize treatments. In this article, we aim to provide a broad audience with a framework for understanding and applying basic principles of learning algorithms.

## What Is Learning?

Most people have not seen every dog that exists, but they can identify a dog when they see it, even if they have never seen that exact dog before. They use pattern recognition by identifying features that are common among most dogs. For example, dogs typically have four legs and a tail, they are furry, and they are often located near a person. People can also identify other animals as “not dogs.” For example, giraffes’ long necks, tall height, and characteristic spot patterns are features that would be quite atypical for dogs. Pattern recognition relies on identifying characteristics typical of a dog, as well as characteristics atypical of a dog.

Although people are good at identifying dogs in many situations based on a combination of features, the average person might not be accurate at distinguishing dogs from

giraffes based only on smell, or only on the appearance of the ears, or only on fur texture. However, people might be able to learn to do so if they had enough exposure or training to recognize patterns in these traits.

## What Is an Algorithm?

An algorithm is a set of instructions or rules that accomplish a task (e.g., categorization).

Example: If a shape has four sides then categorize it as a Quadrilateral; otherwise, categorize it as a Non-Quadrilateral.

## Structure of a Learning Algorithm

A learning algorithm categorizes input data; the algorithm’s ability to accurately categorize the input data improves as it analyzes more data (– **Table 1**). A learning algorithm requires (1) input data, (2) parameters for categorization, and (3) a method of assessing its success (e.g., least squares), allowing for optimization of categorization.

A learning algorithm has the potential to identify complex patterns that are difficult or impossible for humans to distinguish.

Example: Compare handwriting samples from lefties and righties. Produce a best-fit equation that categorizes handwriting as lefty or righty based on letter slant and ink smudge.

Although algorithms are employed frequently in current ophthalmic imaging, the majority of these are static algorithms rather than learning algorithms. One example is the quantification and graphical representation of retinal nerve fiber layer (RNFL) thickness that is produced after RNFL optical coherence tomography (OCT). The RNFL algorithm analyzes

Address for correspondence  
Ariana M. Levin, MD, Department  
of Ophthalmology, University of  
Utah/John A. Moran Eye Center,  
65 Mario Capecchi Drive, Salt Lake  
City, UT 84132  
(e-mail: ariana.levin@hsc.utah.edu).

DOI <https://doi.org/10.1055/s-0039-1685535>.  
ISSN 2475-4757.

Copyright © 2019 by Thieme Medical  
Publishers, Inc., 333 Seventh Avenue,  
New York, NY 10001, USA.  
Tel: +1(212) 584-4662.

License terms



**Table 1** Key terms and definitions

<b>Learning algorithm:</b> an algorithm that adapts as it analyzes more data
<b>Supervised learning:</b> the learning algorithm is trained on input data that are already categorized
<b>Unsupervised learning:</b> input data are not categorized prior to analysis by the learning algorithm
<b>Feature extraction:</b> representation of data by a set of characteristic parameters
<b>Deep learning:</b> complex, multilayered networks of features that relate input data to output data
<b>Cluster analysis:</b> data are categorized into distinct clusters; one cluster represents shared features among the data in that cluster
<b>Archetypal analysis:</b> data are categorized based on patterns common among groups of data
<b>The “Black Box” problem:</b> the process that an algorithm uses to relate input data to output data might be unclear or unavailable to human operators

these images more efficiently than a human and analyzes large amounts of data. However, the algorithm does not change or improve as it analyzes more data. The RNFL algorithm might identify details of an image that are too subtle to be appreciated by a human (e.g., small decreases in thickness), but it cannot identify new patterns (e.g., rates of thickness change that correlate with intraocular pressure changes).

## Input Data: Supervised versus Unsupervised Learning

In supervised learning, input data that are already categorized are used to train the algorithm (e.g., normal macula OCT versus OCT with known diabetic macular edema). Then, the algorithm can be applied to data to which the algorithm is naive. In unsupervised learning, input data are not categorized prior to analysis (e.g., macula OCT from all diabetic patients).

## Categorization of Data: Feature Extraction and Deep Learning

“Feature extraction” or “feature engineering” are processes that reduce input data to a set of representative features that are used to categorize the data. The choice of features is an important step in the process; algorithms that are designed to mimic the human classification experience are usually built around features that are intuitive to human decision-making. For example, vessel branching, tortuosity, and diameter can be represented mathematically and used to categorize input data for an algorithm designed to identify abnormal retinal vasculature.

In deep learning algorithms, complex and layered networks that relate input data to output data are used to identify low-level features in the data which might be too subtle to be identified during the human classification experience.

## Methods of Optimization

In supervised learning, an optimized algorithm maximizes differences between predetermined categories as well as similarities between data within the same category. For example, Kozak et al described an algorithm based on visual field patterns in patients with or without HIV (human immunodeficiency virus).<sup>4</sup> Via training on categorized images, their algorithm “learned” that visual fields of patients with low CD4 (cluster of differentiation 4) tend to have superior field deficits near the blind spot; visual fields of patients without HIV tend not to have these deficits.

Cluster analysis (a tool for unsupervised learning) categorizes data into separate, distinct clusters. Sample et al used cluster analysis to sort uncategorized visual fields into five clusters based on the severity of field loss and the pattern of field loss.<sup>5</sup> Images placed in cluster 1 demonstrated mild, diffuse field loss; cluster 2 demonstrated superior hemifield loss; cluster 3 demonstrated inferior hemifield loss; cluster 4 demonstrated loss of both hemifields. Images in cluster 5 were normal or nearly normal fields.

Archetypal analysis (another tool for unsupervised learning) categorizes data based on common patterns in the data that are not mutually exclusive; one unit of data might fall into multiple categories. Elze et al used archetypal analysis to identify 17 patterns common in glaucomatous visual fields.<sup>6</sup> Some visual fields contained more than one of these common patterns, such as inferior hemifield loss and superior hemifield loss.

## Advantages and Challenges in Learning Algorithm Development

Machine learning algorithms have advantages over human experts in that they are accurate with their calculations, they do not fatigue, and they are reproducible. However, despite the great potential of these systems, they are limited in their scope of only being able to answer the questions that they are designed to answer. Machines might categorize data based on confounding factors that are not appropriate for the clinical question (e.g., cataracts might confound patterns in glaucomatous field measurements). An algorithm’s learning process might be mathematically sound but clinically wrong.

The popular term “black box” as a descriptor of machine learning represents the concern that human operators might not understand how an algorithm is categorizing input to output. For example, consider an algorithm that categorizes fundus photos as diabetic or nondiabetic. The hypothetical algorithm is trained on images that were first graded by grader A (who tends to overcall diabetic eye disease) and grader B (who tends to underdiagnose). If the graders write their initials in the corner of each image that they grade, then the algorithm might learn to categorize images with the letter “A” in the corner as diabetic and images with the letter “B” as nondiabetic. This inappropriate step in the algorithm might be too subtle in the overall

complexity of the algorithm for a human operator to catch. The error (mathematically correct but clinically inappropriate) is hidden in the “black box.”

It is both an advantage and a disadvantage that large datasets are necessary to build useful learning algorithms: a person cannot learn to identify dogs based on fur texture after studying just a few fur samples. An expert human might be able to identify dogs based on fur texture after studying thousands of fur patches. A well-trained learning algorithm could likely achieve the desired categorization more efficiently, requiring less time and fewer data. However, a large enough dataset must be available for the algorithm to learn sufficiently to be useful.

### Application of Concepts: Examples from Published Studies

Gulshan et al published results from a supervised learning algorithm that detects diabetic retinopathy.<sup>7</sup> The input data used to train their learning algorithm were 128,175 retinal images which were first categorized by ophthalmologists. Parameters for categorizing retinal images were defined based on pixel intensity. A neural network was used to optimize categorization with increasing numbers of images. Thus, the ability of their algorithm to detect retinopathy improved as the number of images that it analyzed increased.

Elze et al published results from an unsupervised learning algorithm that classifies subtypes of glaucomatous visual field abnormalities.<sup>6</sup> The authors contrast this method of classification, which relied on no prior knowledge of patterns in glaucomatous fields (because it is unsupervised), with the classification reported in the ocular hypertension treatment study which relied entirely on clinicians' a priori knowledge.<sup>8</sup> Input data included 13,231 visual field measurements from patients with glaucoma or suspicion for glaucoma. Parameters for categorizing images were defined based on location and amplitude of field deviation. Archetypal analysis was used to identify common features in the images (e.g., circular arcuates, partial arcuates, and total loss); it did not separate images into non-overlapping clusters. Each visual field measurement had components of multiple archetypes. As a result of this analysis, the study identified 17 archetypes of glaucomatous visual field loss.

## Conclusions

We presented the foundations of learning algorithms with examples of applications in ophthalmology. The applications are broad, including image interpretation, diagnosis, genetic analysis, risk factor stratification, and personalization of treatments. These systems are best suited to problems that are intellectually difficult for humans to answer (due to complexity of calculations, multidimensionality, etc.) but could be solved by a well-designed computer algorithm.

### Funding

Unrestricted grants from Research to Prevent Blindness and the Utah chapter of the Achievement Rewards for College Scientists Foundation.

### Conflict of Interest

None declared.

## References

1. Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *Nature Digit Med* 2018;1(01):39
2. Sample PA, Goldbaum MH, Chan K, et al. Using machine learning classifiers to identify glaucomatous change earlier in standard visual fields. *Invest Ophthalmol Vis Sci* 2002;43(08):2660–2665
3. Rojas J, Fernandez I, Pastor JC, et al. Development of predictive models of proliferative vitreoretinopathy based on genetic variables: the Retina 4 project. *Invest Ophthalmol Vis Sci* 2009;50(05):2384–2390
4. Goldbaum MH, Kozak I, Hao J, et al. Pattern recognition can detect subtle field defects in eyes of HIV individuals without retinitis under HAART. *Graefes Arch Clin Exp Ophthalmol* 2011;294(04):491–498
5. Sample PA, Chan K, Boden C, et al. Using unsupervised learning with variational bayesian mixture of factor analysis to identify patterns of glaucomatous visual field defects. *Invest Ophthalmol Vis Sci* 2004;45(08):2596–2605
6. Elze T, Pasquale LR, Shen LQ, Chen TC, Wiggs JL, Bex PJ. Patterns of functional vision loss in glaucoma determined with archetypal analysis. *J R Soc Interface* 2015;12(103):20141118
7. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016;316(22):2402–2410
8. Kass MA, Heuer DK, Higginbotham EJ, et al. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. *Arch Ophthalmol* 2002;120(06):701–713, discussion 829–830