Orphanet Journal of Rare Diseases

Clinical Use of Machine Learning in Rare Endocrine Disorders

--Manuscript Draft--

Manuscript Number:	
Full Title:	Clinical Use of Machine Learning in Rare Endocrine Disorders
Article Type:	Review
Section/Category:	Rare endocrinological diseases
Funding Information:	University of Michigan (Lipodystrophy Fund) — Dr Elif A. Oral
Abstract:	Background Machine learning is an area of study in which computers generate solutions to problems they have not been explicitly programmed to solve. The accelerating pace of developments in machine learning has sparked research into its use in medicine. Findings Previous studies have been able to achieve promising results through machine learning in the areas of patient identification, diagnoses, prognoses, and predicting patient and treatment outcomes. These results make machine learning a worthwhile tool to investigate for its use in the identification, diagnosis, and treatment of rare endocrine diseases. In the field of rare endocrine diseases, machine learning presents new possibilities and potential solutions in many different contexts. Recent studies involving machine learning have shown promise in combating under-diagnosis and aiding in identifying patients who may have rare endocrine diseases. Additionally, machine learning has been successful in predicting patient and treatment outcomes. Machine learning also has the capability to learn from images, a capability that has been shown to be effective in diagnosing certain rare diseases. We also address the issue of data scarcity in rare endocrine diseases, consider the obstacles to be overcome for the use of machine learning in this area and explore possible solutions. Finally, we discuss the uses of machine learning with regard to lipodystrophy, a heterogenous cluster of diseases that present challenges for diagnosis that machine learning may be suited to solve. Conclusions The advanced data processing capabilities of machine learning algorithms show great promise in the field of rare endocrine disease research. The personalized and large-scoped analysis machine learning algorithms provide allow for novel identification, diagnosis, and risk prediction methods that may be used to improve patient care and facilitate research on rare endocrine diseases.
Corresponding Author:	Elif A. Oral University of Michigan UNITED STATES
Corresponding Author E-Mail:	eliforal@med.umich.edu
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Michigan
Corresponding Author's Secondary Institution:	
First Author:	Efe Yagiz Akinci, -
First Author Secondary Information:	
Order of Authors:	Efe Yagiz Akinci, - Elif A. Oral
Order of Authors Secondary Information:	

Additional Information:	
Question	**Response**
Is this study a clinical trial?<hr>A clinical trial is defined by the World Health Organisation as 'any research study that prospectively assigns human participants or groups of humans to one or more health-related interventions to evaluate the effects on health outcomes'.	No

Clinical Use of Machine Learning in Rare Endocrine Disorders

Efe Y. Akinci[1], Elif A. Oral[1].

[1] Division of Metabolism, Endocrinology and Diabetes,
Department of Internal Medicine,
University of Michigan Medical School,
Ann Arbor, MI, USA.

Short Title (max 50 characters):

Machine Learning in Rare Endocrine Disorders

Word Count: 4160
Tables: 1
Figures: 3

Address correspondence to:

Elif Arioglu Oral, MD
Professor of Medicine; Metabolism, Endocrinology and Diabetes Division
Department of Internal Medicine
Caswell Diabetes Institute
North Campus Research Complex, 25-3696
University of Michigan, Ann Arbor, Michigan 48109-2800
Phone: 734-615-7271
E-mail: eliforal@med.umich.edu

Abstract

Background

Machine learning is an area of study in which computers generate solutions to problems they have not been explicitly programmed to solve. The accelerating pace of developments in machine learning has sparked research into its use in medicine.

Findings

Previous studies have been able to achieve promising results through machine learning in the areas of patient identification, diagnoses, prognoses, and predicting patient and treatment outcomes. These results make machine learning a worthwhile tool to investigate for its use in the identification, diagnosis, and treatment of rare endocrine diseases. In the field of rare endocrine diseases, machine learning presents new possibilities and potential solutions in many different contexts. Recent studies involving machine learning have shown promise in combating under-diagnosis and aiding in identifying patients who may have rare endocrine diseases. Additionally, machine learning has been successful in predicting patient and treatment outcomes. Machine learning also has the capability to learn from images, a capability that has been shown to be effective in diagnosing certain rare diseases. We also address the issue of data scarcity in rare endocrine diseases, consider the obstacles to be overcome for the use of machine learning in this area and explore possible solutions. Finally, we discuss the uses of machine learning with regard to lipodystrophy, a heterogenous cluster of diseases that present challenges for diagnosis that machine learning may be suited to solve.

Conclusions

The advanced data processing capabilities of machine learning algorithms show great promise in the field of rare endocrine disease research. The personalized and large-scoped analysis

machine learning algorithms provide allow for novel identification, diagnosis, and risk prediction methods that may be used to improve patient care and facilitate research on rare endocrine diseases.

Keywords: Artificial intelligence, machine learning, computer assisted diagnosis, rare endocrine diseases.

List of Abbreviations

CNN – Convolutional Neural Network

TPR – True positive rate

TNR – True negative rate

TN – True negatives

FP – False positives

AUC – Area under the receiver operating characteristic curve

FPR – False positive rate

ASTRAL – Acute Stroke Registry and Analysis of Lausanne

ALP – Alkaline phosphatase

PLP – Pyridoxal 5′-phosphate

AHP – Acute hepatic porphyria

lincRNAs – Long intergenic noncoding RNAs

TSC – Tuberous sclerosis complex

FAF – Fundus autofluorescence imaging

DIM – Deep InfoMax

LD – Lipodystrophy

EHRs – Electronic health records

DXA – Dual-energy X-ray absorptiometry

Background

The technique of machine learning, a process in which a computer generates a solution to a problem through data analysis methods instead of being explicitly programmed to solve the problem, has seen increasing use in the field of medicine. Recent results from studies attempting to use machine learning in medicine have been promising. In particular, machine learning has shown great promise in the areas of diagnosis through medical records and imaging. In this article, we aim to review the current literature on the topic of the use of machine learning in diagnosing rare endocrine diseases and explore possible avenues for further research.

Definition of Machine Learning

Machine learning can be defined as “a computational process that uses input data to achieve a desired task without being literally programmed (i.e., “hard coded”) to produce a particular outcome” (1). In the studies we will be focusing on throughout this review, many different types of machine learning algorithms are widely used. Therefore, we believe that it is important to include a short summary of the terms we will be using to describe and evaluate machine learning methods.

Common Machine Learning Algorithms

Throughout this article, we will review multiple studies that largely use the same few machine learning algorithms. We have summarized some common machine learning concepts in Table 1. Additionally, we believe it is important to provide a brief summary of a few machine learning algorithms in this section.

Support Vector Machines

A support vector machine aims to create a hyperplane or a set of hyperplanes (a plane that is one dimension smaller than the space the data are in, e.g., a line in 2D space, a 2D plane in 3D space) that are able to separate the data into distinct classes (Fig. 1). Support vector machines can use kernels to project features into higher dimensions in cases in which it isn't possible for a hyperplane to linearly classify the data (Fig. 2).

Neural Networks

Neural networks use nodes called "neurons" in a system that aims to mimic the structure of a biological brain. These nodes can take an input, and transform that input based on a series of weights and activation functions to create a new output. This output can then be used for classification and regression purposes. Neural networks often have multiple layers of neurons between the input and output layers. In supervised neural networks, the weights that determine the activation and output of the neurons are automatically adjusted over time based on an error, or cost, function that provides a quantitative value to measure the error of the network. A diagram of a simple neural network is shown in Fig. 3.

Random Forest

A random forest algorithm is a decision-tree based algorithm that uses many uncorrelated decision trees who all "vote" on the classification. Random forest algorithms are useful in cases where there are many input features or overfitting is a concern. Due to the reliance of the algorithm on uncorrelated decision trees, the features must be uncorrelated. Correlated features increase the error rate of the algorithm.

Logistic Regression

Logistic regression fits data points onto a logistic function. This mapping results in a value for each input that we can classify into either class based on a threshold value. Logistic regression is especially useful when considering probability as the values outputted by logistic regression can be interpreted as probabilities.

Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are often used in image analyses. It would be very computationally expensive to use a regular fully connected neural network to analyze images; a regular, 1080p color image would provide $1920 \times 1080 \times 3$ features to analyze when flattened into a regular vector. Instead, CNNs use an approach that is based on various kernels (also called filters) to detect patterns within an image. Pooling layers are then able to extract important features from the previous layer while shrinking the number of features. Multiple convolutional and pooling layers to identify more complex patterns are often used. Finally, the results of the convolutional and pooling layers can be used to classify an object.

<u>Evaluating Machine Learning Performance</u>

Sensitivity, specificity, and the area under the receiver operating characteristic curve are commonly used to evaluate machine learning performance. The sensitivity of a model, also known as the true positive rate (TPR), describes the percentage of positive cases a model is able to identify. This rate is given by the total number of true positives divided by the sum of total true positives and false negatives (FN).

The specificity of a model, also known as the true negative rate (TNR) is given by the total number of true negatives (TN) divided by the sum of true negatives and false positives (FP).

These two metrics are inversely related. As we reduce our threshold to make our model more sensitive, we are bound to include more false positives in our more sensitive classification. Conversely, when we increase our threshold, we are bound to miss some borderline cases, leading us to include more false negatives in our less sensitive classification.

One additional metric we use is the area under the receiver operating characteristic curve (AUC). For this metric, we use the true positive rate and the false positive rate (FPR). TPR and FPR are positively correlated — as we decrease our threshold and our program gets more sensitive, the number of true positives we find will increase, but so will the number of false positives. A ROC curve graphs the TPR and FPR against each other at all possible thresholds. The AUC is a measure of the area under the ROC curve. An ideal prediction algorithm (an algorithm that would classify correctly every time) would have an AUC of 1.0, while an algorithm that guesses randomly would have an AUC of 0.5.

Machine Learning in Medicine

The rapid expansion of computing power and the increase in accessibility to data-processing utilities have led to an increase in the usage of machine learning for various purposes in the field of medicine. For example, personalized medicine has been a strength of machine learning. Machine learning can create models that can easily be used to provide personalized interpretations of data, assist in prognoses, and predict treatment outcomes (2-4).

Machine learning is a promising tool in assisting in diagnoses. Its ability to analyze and classify data, including images, to find complex patterns allows it to accurately draw conclusions from large amounts of data. For example, a recent study (5) utilizing logistic regression and support vector machines, two different methods that incorporate machine learning, showed success in assisting in the diagnosis of transition zone prostate cancer. Another study (6) was able to utilize three different methods of machine learning to identify breast cancer from characteristics of cell nuclei. Furthermore, machine learning has shown promise in predicting disease and

treatment outcomes. In this task, machine learning has often proven itself to be better suited than conventional methods. For example, a previous study (7) aimed to utilize various machine learning algorithms in predicting ischemic stroke outcomes. The machine learning algorithms were trained on 1744 data points using 38 variables such as patient demographics, initial National Institutes of Health Stroke Scale scores, and time from onset to admission. The machine learning algorithms were then compared to a baseline conventional method, the Acute Stroke Registry and Analysis of Lausanne (ASTRAL) score. Using the ASTRAL score, the study achieved an AUC value of 0.839. Although this did not differ significantly from a random forest and logistic regression models (AUC values of 0.857 and 0.849 respectively), a deep neural network was able to achieve an AUC value of 0.888, significantly higher than the conventional ASTRAL score. Notably, when the machine learning algorithms were provided solely with the criteria the ASTRAL system used, the performance of all of the machine learning algorithms was comparable with the ASTRAL score, highlighting the capability of deep learning models to find complex relationships between features.

These capabilities of machine learning show promise for its use in rare endocrine diseases. In an area in which the obscurity and underdiagnosis of diseases present large obstacles, the ability of machine learning to process large amounts of data without being limited by human memory and analytic capabilities can be invaluable.

Machine Learning in Rare Endocrine Diseases

Recently, machine learning has been shown to be very promising in the area of rare disease research. Common problems such as underdiagnosis and misdiagnosis due to obscurity, and misclassification are issues that the large data-processing capabilities of machine learning can be tailored to solve.

Identification of Patients with Rare Diseases

The difficulty of identifying patients with rare diseases continues to be a large obstacle in rare disease research. According to a survey in 2013, the average patient with a rare disease visits 8 doctors and receives 2 to 3 misdiagnoses over the course of 5-8 years before being correctly diagnosed (8). The difficulty in diagnosing rare diseases, in combination with their low prevalence rates, makes them challenging to study. Even the most experienced clinicians may not have encountered a given rare disease in their careers. It is simply not possible for any single clinician to be able to easily diagnose a majority of rare diseases. Therefore, when a patient presents with a rare disease, it is possible, if not likely, that a clinician will not consider the rare disease until they have exhausted more familiar diseases and conditions, even if the information available strongly aligns with the characteristics of the rare disease at hand (9). Machine learning, however, has the capacity to avoid this problem. Once models to identify and diagnose rare diseases have been built, computers can use multiple models to predict whether the patient is likely to have any of the tested diseases at once without obscurity being a factor in how strongly each condition is considered. Once a list of probabilities for possible conditions has been compiled, the clinician can then use their expertise to pinpoint a specific condition and begin appropriate treatment. Therefore, models that can identify and diagnose diseases can provide a substantial improvement in rare disease related patient care.

For example, hypophosphatasia, a rare disease caused by mutations in the *ALPL* gene that encodes tissue nonspecific alkaline phosphatase causing loss-of-function in the enzyme, is a suitable candidate for a relatively simple machine learning model. Garcia-Carretero et al. (10) utilized the biomarkers alkaline phosphatase (ALP) and pyridoxal 5′-phosphate (PLP) to construct two machine learning models to identify the disease. Using a support vector machine that utilized both of the biomarkers, the study was able to achieve an AUC value of 0.936 in identifying hypophosphatasia.

Additionally, improvements in analyzing “big data,” datasets too large to be analyzed by traditional analysis methods, mean that machine learning models can take advantage of feature sets so large that manual analysis could not feasibly find the same complex relationships between variables machine learning models could. One promising use of this capability of machine learning in rare disease research is the analysis of health records. Electronic health records of patients with rare diseases may uncover previously unknown relationships between the disease and various health parameters. Additionally, these relationships could then be used to identify other patients who may have the same rare disease. Underdiagnosis is a large obstacle in rare disease research and treatment, and the powerful data analysis capabilities of machine learning could provide a solution. Indeed, machine learning on electronic health records has previously been successfully used to predict different medical events such as in-hospital mortality, complications, and discharge diagnoses and to analyze patient outcomes (11-13).

In the context of endocrine diseases, previous studies also show promise (14, 15). For example, Cohen et al. (16) were able to train machine learning algorithms on electronic health records provided by the Oregon Health & Science University to identify a rare condition known as acute hepatic porphyria (AHP). The study was able to identify four previously undiagnosed cases of AHP and identified an additional 18 patients who could benefit from testing.

We believe that the results of the aforementioned studies and ongoing efforts may help remedy certain difficulties in the identification of rare diseases in individuals. In cases where the lack of identification is due to the obscurity of a rare disease rather than a lack of data, we believe that scanning of electronic health records using machine learning algorithms may help to locate and treat patients with rare diseases.

Aiding in Diagnosis

Machine learning has shown great strength in aiding diagnoses. In the context of rare endocrine diseases where obscurity and complexity present great challenges, the capability of computers to easily recall large amounts of data makes them uniquely suited to diagnose rare endocrine diseases. Indeed, using computerized decision support systems have previously shown success in assisting to diagnose rare diseases (17). Additionally, the ability of machine learning to easily use a large number of features to estimate outcomes has led to it often outperforming conventional tests and prediction tools without being more invasive (18).

One study (19) used machine learning to identify long intergenic noncoding RNAs (lincRNAs) that may serve as biomarkers in diagnosing pheochromocytomas and paragangliomas. However, in this context, some problems related to the nature of rare diseases may emerge. Similar to the first example, Wallace et al. (20) also utilized machine learning in diagnosing these rare neuroendocrine tumors. The study utilized linear discriminant analysis and metabolic profiling on 186 samples to identify paragangliomas, achieving a sensitivity/specificity ratio of 93.2%/99.2% with an AUC of 0.982. In comparison, evaluating the succinate/fumarate ratio using conventional means yielded a sensitivity/specificity ratio of 88.1%/99.2% with an AUC of 0.96. However, a second model that included ten instead of four metabolites and used formalin-fixed and/or paraffin-embedded tissue in addition to freshly frozen tissue yielded a slightly different result. The study noted that the discrepancy may be explained by metabolite levels differing due to stromal contaminants in the tissue used. In the context of rare diseases, standardization of data is a relevant discussion, as data are often collected from various sources. The lack of standardization that may result from such circumstances may pose an obstacle in training machine learning models on rare diseases.

Machine learning has also proven to be useful in differentiating rare diseases. Arlt et al. (21) attempted to differentiate adrenocortical carcinomas and adrenocortical adenomas. The study

used steroid excretion analysis by mass spectrometry to obtain measurements on certain steroid markers. The study then used generalized matrix learning vector quantization and trained the model using 32 steroid markers as features. The model was able to achieve an AUC value of 0.965 in differentiating the two conditions.

Furthermore, unsupervised machine learning algorithms provide a promising approach to tasks such as gene clustering. Unsupervised machine learning algorithms are suited to discover complex polygenic attributes and identify genes and other biomarkers of interest without any manual labeling. Machine learning algorithms have previously been used in this context to gain insight into various conditions (22-24). This approach may uncover new avenues of exploration for research on causes and treatment methods for rare diseases.

Aiding in Prognosis and Risk Prediction

A future in which patients are able to easily check their risk status for different diseases would allow them to proactively make healthier decisions, seek earlier treatment, and be more observant of symptoms for diseases they are at risk for. Machine learning has previously been successfully used in endocrine diseases to predict different patient risks and treatment outcomes. For example, machine learning algorithms have successfully been able to assess risk for diabetes (25-27) and its potential complications (28), evaluate treatment outcomes for various endocrine diseases (29-32), and predict risk for adverse events and complications (33-35). Additionally, machine learning has been used to identify new risk factors for rare endocrine diseases (36), and predict treatment outcomes in other less common endocrine diseases such as Cushing's disease (37).

Furthermore, biobanks, repositories that store health information that can be used to retroactively analyze health histories and outcomes, have seen use in estimating underdiagnosis and risk prediction (38-40). Through its capability to establish complex

relationships in large datasets and various biobanks, machine learning presents itself as a suitable tool to aid in prognoses and to establish patient risk for rare endocrine diseases.

Machine Learning in Medical Imaging and Scarcity of Data

A future in which clinicians could get second opinions on various imaging procedures at the click of a button or a future in which they could analyze images for diseases they may have never heard of would serve to reduce the effects that misdiagnoses and underdiagnoses of rare disease cause. Recent machine learning research and the exponential increase in graphics processing power have made convolutional neural networks (CNNs) a very strong candidate for this task. CNNs work by identifying patterns in images that are related to the outcome. This ability of CNNs makes them suited to analyze medical images to aid in the identification and diagnosis of rare diseases. Imaging-based machine learning models have already been extensively used in various medical fields (41-45).

Although examples in conventional diseases achieve very impressive results, when considering CNN-based methods in the context of rare endocrine diseases, we encounter a problem. For example, although Esteva, et al. (46) were able to achieve excellent results in classifying different types of skin cancer through clinical imaging, the study used 129,450 images to train its machine learning algorithm. CNNs often require large datasets of images to train on, which presents an obstacle for their use in rare disease research. However, there are various methods to adapt CNNs to overcome data scarcity issues that may be present in rare disease research. One method is referred to as “data augmentation”. In this method, the training dataset is augmented through image manipulation such as rotating, cropping, or changing the zoom levels of images. A recent study (47) utilized data augmentation techniques to train a CNN to detect tuberous sclerosis complex (TSC) on 138 images (69 TSC 69 control). When tested on an

unaugmented test dataset (50 images 25 TSC 25 control), the CNN was able to achieve a sensitivity/specificity rate of 95%/95% with an AUC of 0.99.

Another machine learning method that can be used to adapt CNNs for rare disease research is transfer learning. Transfer learning is a method in which machine learning models that have been previously trained on other sets of data are adapted into a new model for the task at hand. A study (48) in 2016 attempted to use transfer learning to develop a model to classify lung tissue patterns from CT scans into different interstitial lung diseases. Through transferring layers from a neural network trained on a texture recognition dataset, the study was able to achieve an absolute performance increase of 2%. Similarly, another study (49) that aimed to classify different inherited retinal diseases through fundus autofluorescence imaging (FAF) trained and validated a CNN using 389 FAF images. The study used transfer learning to build a machine learning model that was first trained on the ImageNet dataset, and then fine-tuned on the 389 FAF images to classify them into different groups. The CNN was then tested on a test set of 94 FAF images with an AUC ≥ 0.989 for all tested diseases.

Recent algorithms have also shown promise in learning from fewer images and data points. A recent study (50) attempted to classify retinal diseases from the EyePACS database with the aim of observing performance differences between various machine learning algorithms as the number of samples used to train the models changed. As a baseline, the study used a traditional fine-tuned ResNet algorithm, referred to as RES-FT. The study then compared the baseline to various algorithms that utilized a network model called Augmented Multiscale Deep InfoMax, which is based on the concept of self-supervision. Using 5120 samples, all of the algorithms scored similarly, with an AUC value of 0.8330 for the baseline and 0.8348 for the Deep InfoMax (DIM) algorithm. However, at 160 samples, the baseline saw a dramatic drop to an AUC value of 0.6585, while the DIM algorithm was able to achieve an AUC value of 0.7467. This trend continued for sample sizes of 40 and 10, with the RES-FT algorithm achieving AUC

values of 0.5671 and 0.5178 respectively, while the DIM algorithm achieved AUC values of 0.6760 and 0.5778 respectively.

These studies illustrate multiple methods that can be used to overcome obstacles that may occur when using machine learning models on smaller amounts of data. Additionally, machine learning has shown success in augmenting images to make obtaining data more accessible (51), an ability that may assist in collecting data for the analysis of rare endocrine diseases. Machine learning models could also help in the early identification of diseases through photos that patients could take themselves. This would allow for patients to diagnose any diseases they may have in their early stages, at which point their symptoms may not be severe enough for a visit to a clinician. One implementation of this idea was a neural network that was designed to identify genetic conditions from pictures of faces (52). This network was successfully able to classify genetic conditions using pictures taken without any special equipment. An expansion of this model into rare diseases may prove successful in identifying obscure diseases.

Using these methods, we believe that CNNs continue to show promise in identifying and diagnosing rare endocrine diseases through medical imaging.

Lipodystrophy – A Promising Application of Machine Learning in Rare Diseases

Lipodystrophy (LD) syndromes are characterized by the lack of adipose tissue in the body, causing complications such as insulin resistance, ectopic steatosis, and hyperlipidemia (53). Lipodystrophy syndromes are considered rare diseases. The combination of a low prevalence rate and the difficulty in diagnosing rare diseases such as lipodystrophy makes it exceedingly difficult to conduct studies on this disease, as patients are scarce and often undiagnosed.

As an example, Gonzaga-Jauregui, et al. (53) estimate the prevalence of LD among the population to be near 47.3 cases per 1,000,000 people. In contrast, Chiquette et al. (54) estimate the prevalence of LD among the general population to be 3.07 cases per 1,000,000

people. The variance between these two figures is most likely due to the criteria used when determining whether a patient could be considered a case of LD. Chiquette, et al. (54) used database-dependent inclusion-exclusion criteria to identify cases of LD, while Gonzaga-Jauregui, et al. (53) observed the expected comorbidities associated with LD to achieve the same task.

The significance of the selected criteria when attempting to identify lipodystrophy from health records is clear from the preceding studies. To obtain a more accurate prevalence value of LD in the general population, we propose a machine learning model similar to Zheng, et al. (15). The study was able to accurately identify type-2 diabetes through electronic health records (EHRs) that included previous medical conditions, and we believe that a similar approach may yield promising results in identifying LD. Indeed, it is well known that lipodystrophy is associated with a multitude of notable comorbidities and signs such as insulin resistance and dyslipidemia that healthcare providers may not immediately connect to LD, but that will nevertheless be included in EHRs. Furthermore, as LD is known to be an underdiagnosed disease, a successful machine learning algorithm to identify cases through EHRs will have further implications in research and treatment of the disease.

Further on the topic of underdiagnosis, clinicians may not always have the required experience with LD to accurately diagnose it. A recent study (55) has developed a method called “fat shadows” to obtain a visual representation of adipose tissue distribution across the body from dual-energy X-ray absorptiometry (DXA) images. We propose that a convolutional neural network may be able to identify patterns corresponding to LD from fat shadows, which could then enable the identification of LD without requiring an expert on the disease. Indeed, there have been previously conducted studies (56, 57) that were able to successfully utilize machine learning algorithms to analyze DXA images. A successful machine learning algorithm that could identify LD from images would be invaluable in combatting underdiagnosis in LD studies.

Conclusions

Current literature on machine learning in rare endocrine disorders suggests that machine learning may hold great promise in overcoming major issues in rare disease research. We believe that the diagnostic and identification capabilities of machine learning can combat the misdiagnoses and under-diagnosis that present obstacles for patients seeking care and researchers who wish to study rare endocrine disorders. Additionally, the prognostic capabilities of machine learning can help clinicians better understand the clinical course rare endocrine disorders will take in an individual and decide on the best treatment for a specific case. Finally, we believe that through methods of remedying problems caused by data scarcity, machine learning will continue to play an increasing role in all aspects of rare endocrine disease research in the near and long-term future. It would be important to recognize this potential and train the next generation of clinicians and researchers in the field to be fully equipped to make use of these methods.

References

1. El Naqa I, Murphy MJ. What Is Machine Learning? In: El Naqa I, Li R, Murphy MJ, editors. Machine Learning in Radiation Oncology: Theory and Applications. Cham: Springer International Publishing; 2015. p. 3-11.
2. Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. Curr Cardiol Rep. 2014;16(1):441.
3. Jiang R, Abbott CC, Jiang T, Du Y, Espinoza R, Narr KL, et al. SMRI Biomarkers Predict Electroconvulsive Treatment Outcomes: Accuracy with Independent Data Sets. Neuropsychopharmacology. 2018;43(5):1078-87.
4. Montazeri M, Montazeri M, Montazeri M, Beigzadeh A. Machine learning models in breast cancer survival prediction. Technol Health Care. 2016;24(1):31-42.
5. Wu M, Krishna S, Thornhill RE, Flood TA, McInnes MDF, Schieda N. Transition zone prostate cancer: Logistic regression and machine-learning models of quantitative ADC, shape and texture features are highly accurate for diagnosis. J Magn Reson Imaging. 2019;50(3):940-50.
6. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol. 2019;19(1):64.
7. Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke. Stroke. 2019;50(5):1263-5.
8. Shire. Rare Disease Impact Report 2013 [10]. Available from: https://globalgenes.org/wp-content/uploads/2013/04/ShireReport-1.pdf.
9. Lee CS, Nagy PG, Weaver SJ, Newman-Toker DE. Cognitive and system factors contributing to diagnostic errors in radiology. AJR Am J Roentgenol. 2013;201(3):611-7.
10. Garcia-Carretero R, Olid-Velilla M, Perez-Torrella D, Torres-Pacho N, Darnaude-Ortiz MT, Bustamate-Zuloeta AD, et al. Predictive modeling of hypophosphatasia based on a case series of adult patients with persistent hypophosphatasemia. Osteoporos Int. 2021.
11. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. NPJ Digit Med. 2018;1:18.
12. Miotto R, Li L, Kidd BA, Dudley JT. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. Sci Rep. 2016;6:26094.
13. Ruan Y, Bellot A, Moysova Z, Tan GD, Lumb A, Davies J, et al. Predicting the Risk of Inpatient Hypoglycemia With Machine Learning Using Electronic Health Records. Diabetes Care. 2020;43(7):1504-11.
14. Wei Z, Wang K, Qu HQ, Zhang H, Bradfield J, Kim C, et al. From disease association to risk assessment: an optimistic view from genome-wide association studies on type 1 diabetes. PLoS Genet. 2009;5(10):e1000678.
15. Zheng T, Xie W, Xu L, He X, Zhang Y, You M, et al. A machine learning-based framework to identify type 2 diabetes through electronic health records. Int J Med Inform. 2017;97:120-7.
16. Cohen AM, Chamberlin S, Deloughery T, Nguyen M, Bedrick S, Meninger S, et al. Detecting rare diseases in electronic health records using machine learning and knowledge engineering: Case study of acute hepatic porphyria. PLoS One. 2020;15(7):e0235574.

17. Ronicke S, Hirsch MC, Turk E, Larionov K, Tientcheu D, Wagner AD. Can a decision support system accelerate rare disease diagnosis? Evaluating the potential impact of Ada DX in a retrospective study. Orphanet J Rare Dis. 2019;14(1):69.
18. Atabaki-Pasdar N, Ohlsson M, Vinuela A, Frau F, Pomares-Millan H, Haid M, et al. Predicting and elucidating the etiology of fatty liver disease: A machine learning modeling and validation study in the IMI DIRECT cohorts. PLoS Med. 2020;17(6):e1003149.
19. Ghosal S, Das S, Pang Y, Gonzales MK, Huynh TT, Yang Y, et al. Long intergenic noncoding RNA profiles of pheochromocytoma and paraganglioma: A novel prognostic biomarker. Int J Cancer. 2020;146(8):2326-35.
20. Wallace PW, Conrad C, Bruckmann S, Pang Y, Caleiras E, Murakami M, et al. Metabolomics, machine learning and immunohistochemistry to predict succinate dehydrogenase mutational status in phaeochromocytomas and paragangliomas. J Pathol. 2020;251(4):378-87.
21. Arlt W, Biehl M, Taylor AE, Hahner S, Libe R, Hughes BA, et al. Urine steroid metabolomics as a biomarker tool for detecting malignancy in adrenal tumors. J Clin Endocrinol Metab. 2011;96(12):3775-84.
22. van IDGP, Szuhai K, Briaire-de Bruijn IH, Kostine M, Kuijjer ML, Bovee J. Machine learning analysis of gene expression data reveals novel diagnostic and prognostic biomarkers and identifies therapeutic targets for soft tissue sarcomas. PLoS Comput Biol. 2019;15(2):e1006826.
23. Gagliano L, Assi EB, Toffa DH, Nguyen DK, Sawan M. Unsupervised Clustering of HRV Features Reveals Preictal Changes in Human Epilepsy. Annu Int Conf IEEE Eng Med Biol Soc. 2020;2020:698-701.
24. Laas E, Ballester M, Cortez A, Graesslin O, Darai E. Unsupervised Clustering of Immunohistochemical Markers to Define High-Risk Endometrial Cancer. Pathol Oncol Res. 2019;25(2):461-9.
25. Joglekar MV, Wong WKM, Ema FK, Georgiou HM, Shub A, Hardikar AA, et al. Postpartum circulating microRNA enhances prediction of future type 2 diabetes in women with previous gestational diabetes. Diabetologia. 2021.
26. van der Meer TP, Wolffenbuttel BHR, Patel CJ. Data-driven assessment, contextualisation and implementation of 134 variables in the risk for type 2 diabetes: an analysis of Lifelines, a prospective cohort study in the Netherlands. Diabetologia. 2021.
27. Qiao N, Shen M, He W, He M, Zhang Z, Ye H, et al. Machine learning in predicting early remission in patients after surgical treatment of acromegaly: a multicenter study. Pituitary. 2021;24(1):53-61.
28. Huang J, Huth C, Covic M, Troll M, Adam J, Zukunft S, et al. Machine Learning Approaches Reveal Metabolic Signatures of Incident Chronic Kidney Disease in Individuals With Prediabetes and Type 2 Diabetes. Diabetes. 2020;69(12):2756-65.
29. Park YW, Eom J, Kim S, Kim H, Ahn SS, Ku CR, et al. Radiomics with Ensemble Machine Learning Predicts Dopamine Agonist Response in Patients with Prolactinoma. J Clin Endocrinol Metab. 2021.
30. Wildemberg LE, da Silva Camacho AH, Miranda RL, Elias PCL, de Castro Musolino NR, Nazato D, et al. Machine learning-based prediction model for treatment of acromegaly with first-generation somatostatin receptor ligands. J Clin Endocrinol Metab. 2021.

31. Negro R, Rucco M, Creanza A, Mormile A, Limone PP, Garberoglio R, et al. Machine Learning Prediction of Radiofrequency Thermal Ablation Efficacy: A New Option to Optimize Thyroid Nodule Selection. Eur Thyroid J. 2020;9(4):205-12.
32. Baum A, Scarpa J, Bruzelius E, Tamler R, Basu S, Faghmous J. Targeting weight loss interventions to reduce cardiovascular complications of type 2 diabetes: a machine learning-based post-hoc analysis of heterogeneous treatment effects in the Look AHEAD trial. Lancet Diabetes Endocrinol. 2017;5(10):808-15.
33. Mathioudakis NN, Abusamaan MS, Shakarchi AF, Sokolinsky S, Fayzullin S, McGready J, et al. Development and Validation of a Machine Learning Model to Predict Near-Term Risk of Iatrogenic Hypoglycemia in Hospitalized Patients. JAMA Netw Open. 2021;4(1):e2030913.
34. Voglis S, van Niftrik CHB, Staartjes VE, Brandi G, Tschopp O, Regli L, et al. Feasibility of machine learning based predictive modelling of postoperative hyponatremia after pituitary surgery. Pituitary. 2020;23(5):543-51.
35. Liu Y, Liu X, Hong X, Liu P, Bao X, Yao Y, et al. Prediction of Recurrence after Transsphenoidal Surgery for Cushing's Disease: The Use of Machine Learning Algorithms. Neuroendocrinology. 2019;108(3):201-10.
36. Andrade-Campos MM, de Frutos LL, Cebolla JJ, Serrano-Gonzalo I, Medrano-Engay B, Roca-Espiau M, et al. Identification of risk features for complication in Gaucher's disease patients: a machine learning analysis of the Spanish registry of Gaucher disease. Orphanet J Rare Dis. 2020;15(1):256.
37. Fan Y, Li Y, Bao X, Zhu H, Lu L, Yao Y, et al. Development of Machine Learning Models for Predicting Postoperative Delayed Remission in Patients With Cushing's Disease. J Clin Endocrinol Metab. 2021;106(1):e217-e31.
38. Dickey AK, Quick C, Ducamp S, Zhu Z, Feng YA, Naik H, et al. Evidence in the UK Biobank for the underdiagnosis of erythropoietic protoporphyria. Genet Med. 2021;23(1):140-8.
39. Muller DC, Johansson M, Brennan P. Lung Cancer Risk Prediction Model Incorporating Lung Function: Development and Validation in the UK Biobank Prospective Cohort Study. J Clin Oncol. 2017;35(8):861-9.
40. Alaa AM, Bolton T, Di Angelantonio E, Rudd JHF, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. PLoS One. 2019;14(5):e0213653.
41. Martin-Isla C, Campello VM, Izquierdo C, Raisi-Estabragh Z, Baessler B, Petersen SE, et al. Image-Based Cardiac Diagnosis With Machine Learning: A Review. Front Cardiovasc Med. 2020;7:1.
42. Thomsen K, Iversen L, Titlestad TL, Winther O. Systematic review of machine learning for diagnosis and prognosis in dermatology. J Dermatolog Treat. 2020;31(5):496-510.
43. Yassin NIR, Omran S, El Houby EMF, Allam H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. Comput Methods Programs Biomed. 2018;156:25-45.
44. Williams BM, Borroni D, Liu R, Zhao Y, Zhang J, Lim J, et al. An artificial intelligence-based deep learning algorithm for the diagnosis of diabetic neuropathy using corneal confocal microscopy: a development and validation study. Diabetologia. 2020;63(2):419-30.

45. Wei R, Jiang C, Gao J, Xu P, Zhang D, Sun Z, et al. Deep-Learning Approach to Automatic Identification of Facial Anomalies in Endocrine Disorders. Neuroendocrinology. 2020;110(5):328-37.
46. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-8.
47. Sanchez Fernandez I, Yang E, Calvachi P, Amengual-Gual M, Wu JY, Krueger D, et al. Deep learning in rare disease. Detection of tubers in tuberous sclerosis complex. PLoS One. 2020;15(4):e0232376.
48. Christodoulidis S, Anthimopoulos M, Ebner L, Christe A, Mougiakakou S. Multisource Transfer Learning With Convolutional Neural Networks for Lung Pattern Analysis. IEEE J Biomed Health Inform. 2017;21(1):76-84.
49. Miere A, Le Meur T, Bitton K, Pallone C, Semoun O, Capuano V, et al. Deep Learning-Based Classification of Inherited Retinal Diseases Using Fundus Autofluorescence. J Clin Med. 2020;9(10).
50. Burlina P, Paul W, Mathew P, Joshi N, Pacheco KD, Bressler NM. Low-Shot Deep Learning of Diabetic Retinopathy With Potential Applications to Address Artificial Intelligence Bias in Retinal Diagnostics and Rare Ophthalmic Diseases. JAMA Ophthalmol. 2020;138(10):1070-7.
51. Jin L, Liu B, Zhao F, Hahn S, Dong B, Song R, et al. Deep learning enables structured illumination microscopy with low light levels and enhanced speed. Nat Commun. 2020;11(1):1934.
52. Gurovich Y, Hanani Y, Bar O, Nadav G, Fleischer N, Gelbman D, et al. Identifying facial phenotypes of genetic disorders using deep learning. Nat Med. 2019;25(1):60-4.
53. Gonzaga-Jauregui C, Ge W, Staples J, Van Hout C, Yadav A, Colonie R, et al. Clinical and Molecular Prevalence of Lipodystrophy in an Unascertained Large Clinical Care Cohort. Diabetes. 2020;69(2):249-58.
54. Chiquette E, Oral EA, Garg A, Araujo-Vilar D, Dhankhar P. Estimating the prevalence of generalized and partial lipodystrophy: findings and challenges. Diabetes Metab Syndr Obes. 2017;10:375-83.
55. Meral R, Ryan BJ, Malandrino N, Jalal A, Neidert AH, Muniyappa R, et al. "Fat Shadows" From DXA for the Qualitative Assessment of Lipodystrophy: When a Picture Is Worth a Thousand Numbers. Diabetes Care. 2018;41(10):2255-8.
56. Hussain D, Al-Antari MA, Al-Masni MA, Han SM, Kim TS. Femur segmentation in DXA imaging using a machine learning decision tree. J Xray Sci Technol. 2018;26(5):727-46.
57. Hussain D, Han SM. Computer-aided osteoporosis detection from DXA imaging. Comput Methods Programs Biomed. 2019;173:87-107.

Table

Table 1. A summary of commonly used machine learning algorithm types.

Type of Machine Learning	Description	Uses
Supervised Learning	We provide the computer with a method of measuring the error of its attempt. For every input, we know what output we should get. The computer is then able to analyze its mistakes and adjust its solution in order to achieve greater accuracy through methods such as back-propagation.	Useful in cases where there is easy access to labeled data. • Use of an algorithm to predict treatment outcomes when there is access to a dataset that includes past patient outcomes • Use of an algorithm to diagnose a disease when there is access to health records that differentiate between those who have the disease and those who do not.
Unsupervised Learning	An unsupervised machine learning algorithm trains on datasets where we do not know the labels and do not have a way of evaluating the error of a given attempt. Because we do not have a way of evaluating whether are predictions are "correct," unsupervised machine learning aims to find patterns and structure in data.	Useful when we do not *a priori* want to or know how to label data. • The algorithm can scan the data and identify features and come up with groupings.
Semi-Supervised Learning	A semi-supervised machine learning model is a mixture of supervised and unsupervised learning concepts. To train a semi-supervised machine learning model, we use labeled and unlabeled sets of data. The semi-supervised model is still able to use the unlabeled images to extract certain features that can improve performance.	Useful when it would be too laborious to manually label every image or feature in a dataset. • Users could label some of the images or features that are important, and allow the machine to use the unlabeled images to supplement the labeled images

Legends for Figures

Figure 1. Support vector machines.
Figure 1A shows a randomly scattered array of points that have been classified into two groups based on their positions ($2x < y$). Figure 1B shows the hyperplane a support vector machine has generated to separate the two groups.

Figure 2. Kernel method.
Figure 2-A shows a randomly scattered array of points that have been classified into two groups based on whether they are more than 3.5 units away from the center of the plot (5, 5). Figure 2-B shows a kernel mapping so that a hyperplane ($Z=12.25$) could linearly separate the two groups.

Figure 3. Neural Network
Figure 3 shows a simple neural network with an input layer with six features, an output layer consisting of a single node, and three hidden densely connected layers.

Declarations

Ethics approval and consent to participate

No formal Institutional Review Board (IRB) approval is required as the manuscript reviews published literature.

Consent for publication

Not applicable.

Availability of data and materials

There were no specific datasets analyzed for this review article.

Competing interests

EYA has nothing to disclose. EAO reports the following conflicts: Grant support: Aegerion Pharmaceuticals (now Amryt Pharmaceuticals), Ionis Pharmaceuticals, Akcea Therapeutics, Gemphire Therapeutics, GI Dynamics (current), AstraZeneca (2015-2017). Consultant or Advisor: AstraZeneca, Thera Therapeutics, and BMS (past), Aegerion Pharmaceuticals (now Amryt Pharmaceuticals), Akcea Therapeutics, Ionis Pharmaceuticals, Regeneron Pharmaceuticals (current). Drug support: Aegerion Pharmaceuticals (now Amryt Pharmaceuticals), Akcea Therapeutics, Rhythm Pharmaceuticals (all current). Other support: Aegerion Pharmaceuticals (now Amryt Pharmaceuticals), Regeneron Pharmaceuticals (current).

Funding

The work was partially supported by the Lipodystrophy Fund at the University of Michigan Medical School.

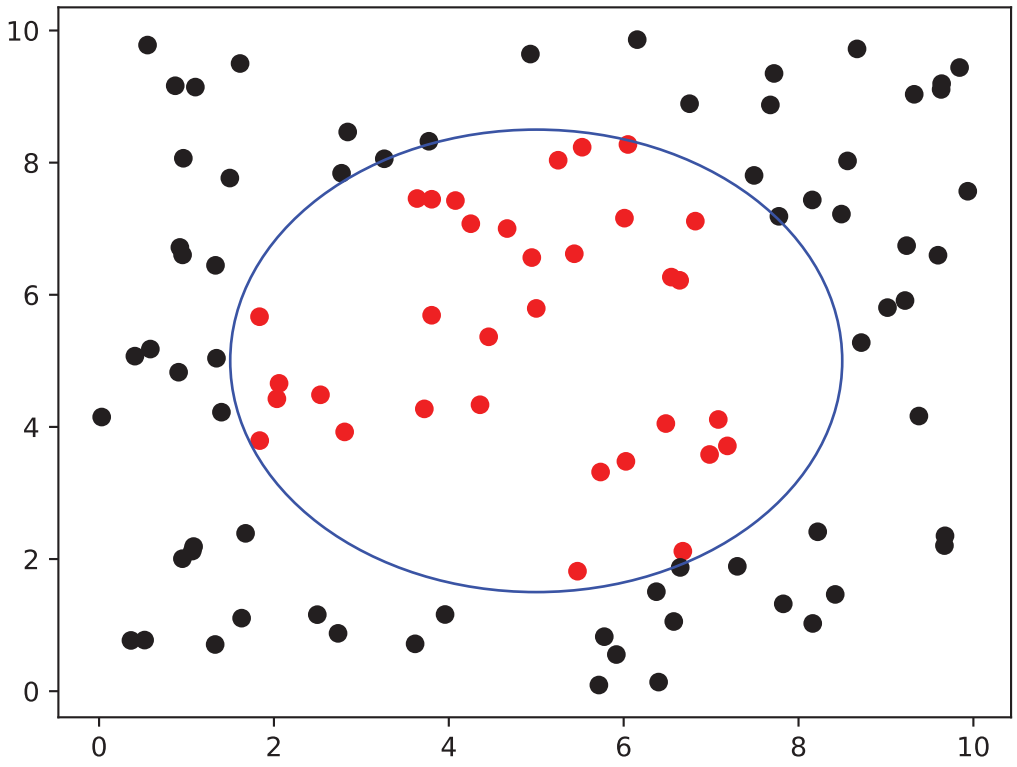
Authors' contributions

EYA gathered the data and wrote the manuscript. EAO reviewed and edited the manuscript. EAO is responsible for the integrity of the data and the conduct of the study. Both authors read and provided critical input on different versions of the manuscript and approved the final version.

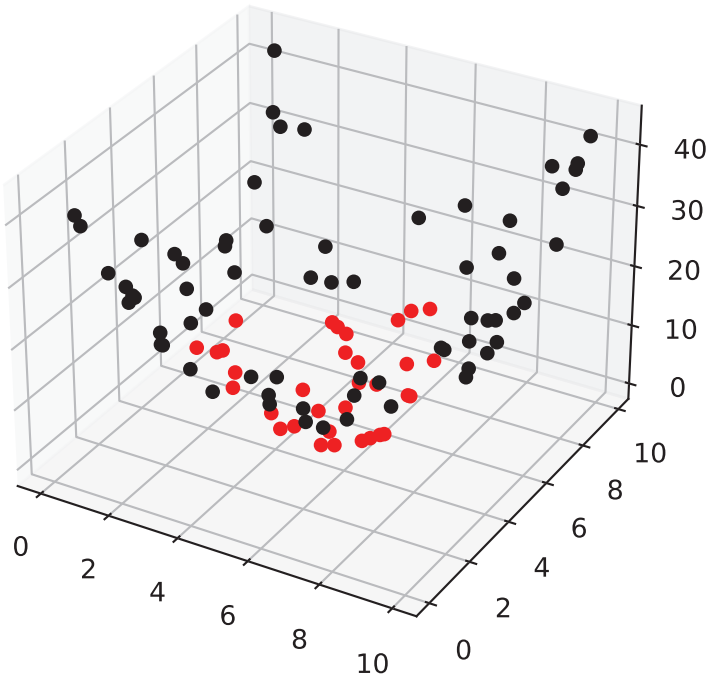
Acknowledgements

We thank Rasimcan Meral, MD for introducing machine learning to our lab.

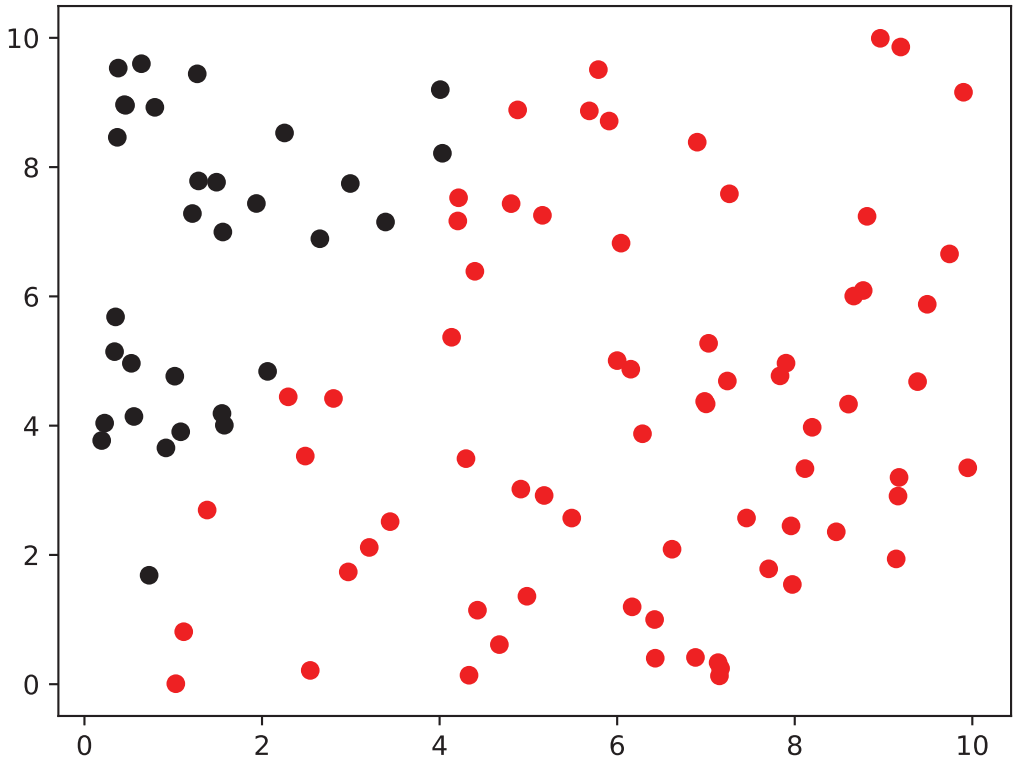
A.



B.



A



B.

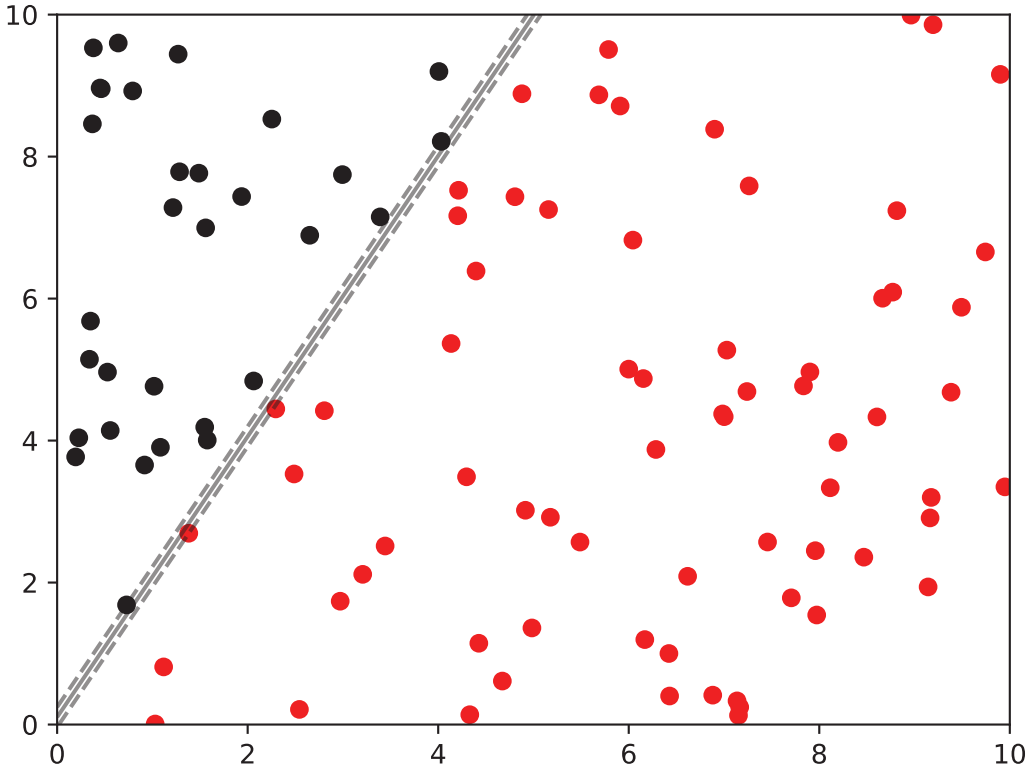
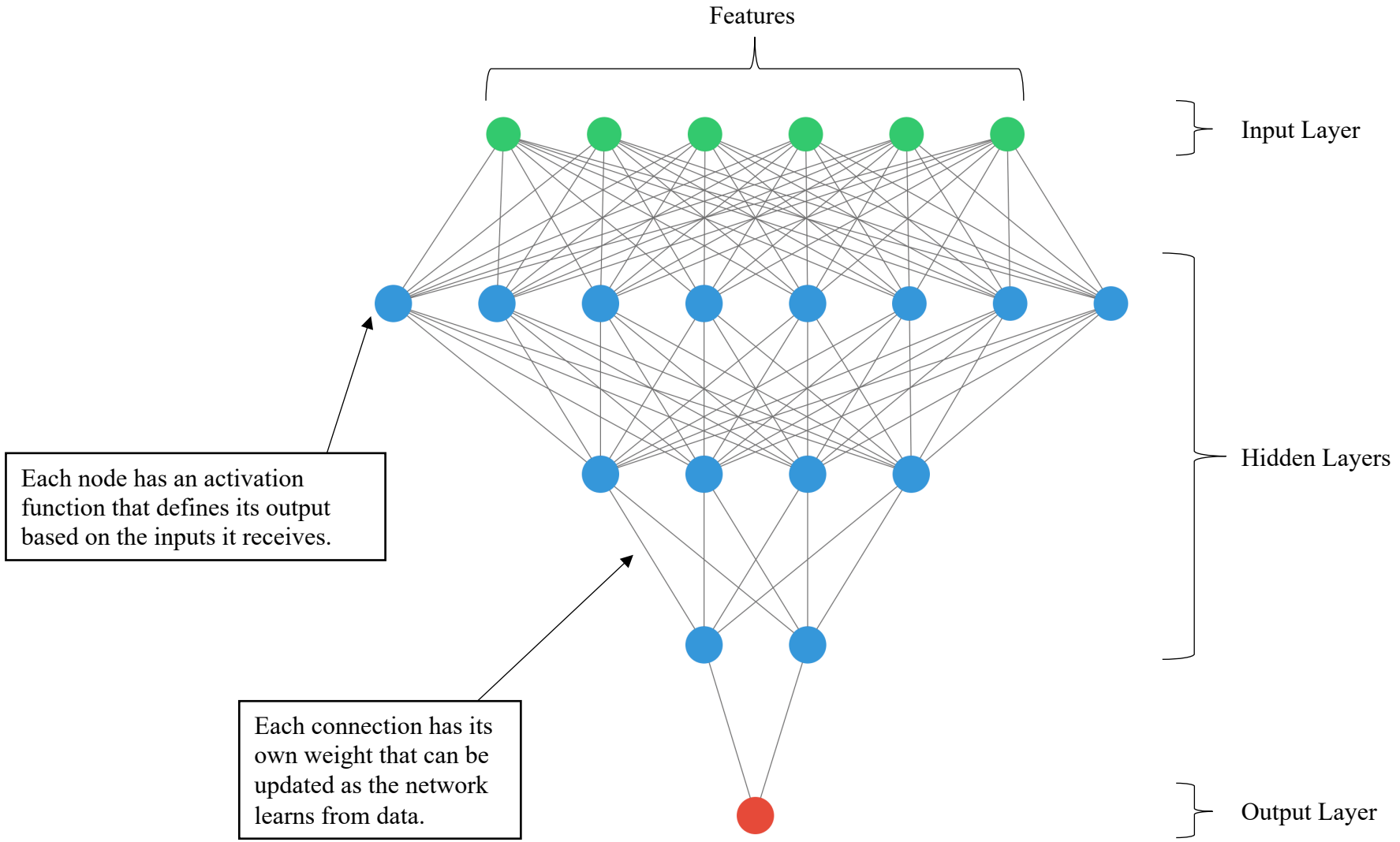


Figure 3



Cover letter