We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Human Active Learning

Akram M. Radwan

Additional information is available at the end of the chapter

Abstract

Active machine learning (AML) is a popular research area in machine learning. It allows selection of the most informative instances in training data of the domain for manual labeling. AML aims to produce a highly accurate classifier using as few labeled instances as possible, thereby minimizing the cost of obtaining labeled data. As machines can learn from experience like humans do, using AML for human category learning may help human learning become more efficient and hence reduce the cost of teaching. This chapter is a review of recent research literature concerning the use of AML technique to enhance human learning and teaching. There are a few studies on the applications of AML to the human category learning domain. The most interesting study was by Castro et al., which showed that humans learn faster with better performance when they can actively select the informative instances from a pool of unlabeled data instead of random sampling. Although AML can facilitate object categorization for humans, there are still many challenges and questions that need to be addressed in the use of AML for modeling human categorization. In this chapter, we will discuss some of these challenges.

Keywords: active learning, machine learning, human learning, cognitive science, categorization

1. Introduction

With the growing amount of data produced daily, the need of techniques to handle these data has been increased. Machine learning (ML) is a prominent area of computer science that evolved from the study of pattern recognition and computational learning theory in artificial intelligence (AI) [1]. ML is algorithms that are able to learn from data, identify patterns in observed data, and build models that make predictions about unseen data. Algorithm or model enables

computer (machine) to learn from data. Over the past decades, learning algorithms have found widespread applications in numerous areas such as computer vision, object recognition, web search, natural language processing, emotion recognition, etc. The performance of different ML algorithms strongly depends on the size and structure of dataset of the domain.

Supervised machine learning learns from the available data (experience), which is given in the form of training data (instances). The knowledge induced from the data can then be used for descriptive or predictive purposes. Supervised learning problems can be categorized as either classification or regression, depending on the output label of the data [2]. Classification problems assign a discrete class label to input instance, while regression problems have continuous numeric values. Classification is a function that assigns a new object (or instance) as belonging to one of the predefined classes [1]. The goal of classification is to accurately predict the target class for each instance in the data.

Figure 1 shows the workflow in supervised learning. We can see that there are two different steps: training and prediction. During training, a feature extractor is used to convert each input (training data instance) to a feature set. These feature sets with labels are fed into the learning algorithm to generate a classifier model. During prediction, the same feature extractor is used to convert unseen inputs to feature sets that are then fed into the model, which generates predicted labels [3, 4].

ML with many other disciplines of AI are gaining popularity, and they have been used in numerous fields and industries, including finance, healthcare, education, and psychology. Since learning is an important aspect of intelligent behavior, ML can be used commonly for data analysis in psychology and cognitive science. Recently, ML methods have been investigated in experimental psychology and human categorization.

In ML, active learning refers to an approach that selects the queries (instances) for labeling from a large pool of unlabeled data [5, 6]. In most cases, an active learning algorithm outperformed random sampling method and reduced number of instances that necessary to achieve similar performance. Active learning is often used for problems where it is difficult (expensive and/or time-consuming) to obtain labeled training data [7, 8].
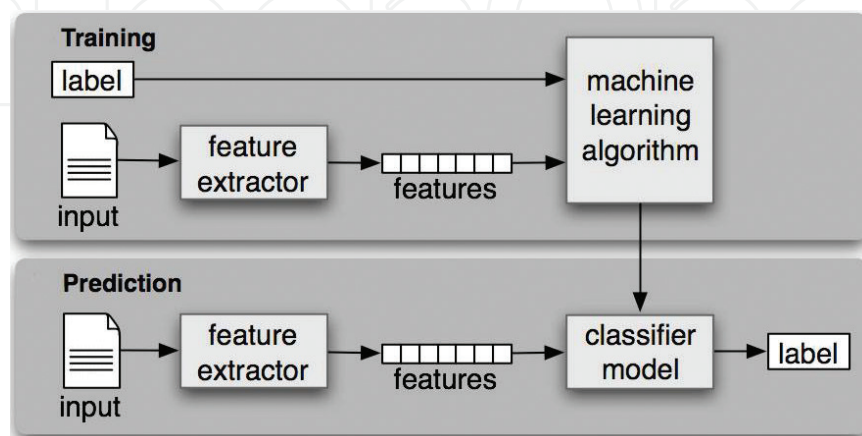

Figure 1. The architecture of a supervised classification.

Researchers have shown that active learning is beneficial in many domains [9], including education [10], machine learning [5], remote sensing [11], and cognitive science [12].

2. Active machine learning

Semi-supervised learning (or SSL) has attracted a highly considerable amount of interest in ML. SSL techniques allow classifiers or learners to learn from labeled and unlabeled data at the same time [13, 14]. Typically, they are used when we have a small-size labeled dataset with a large-size unlabeled dataset. **Figure 2** intuitively shows the difference between supervised and semi-supervised learning. Actually most real-world learning scenarios are SSL. During the last two decades, SSL methods such as active learning, co-training, and co-testing have significantly improved learning performance in various applications.

When a machine learning model is trained, learning is performed on a random subset of all the available sets of labeled training data. We will refer to this mode of learning as passive learning (PL). In PL mode of learning, the classifier (learner) does not participate interactively with the
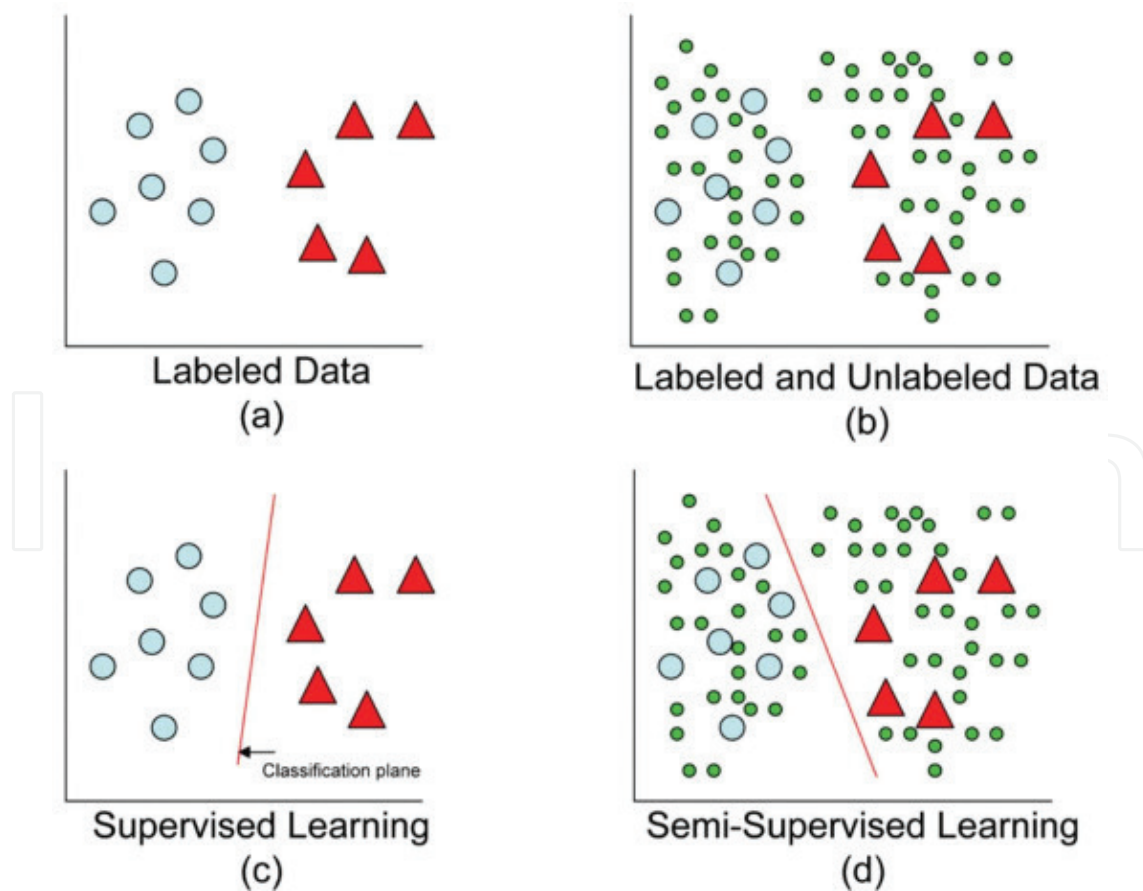

Figure 2. Supervised vs. semi-supervised learning.

teacher [13]. A passive learner receives a random dataset from the world and then produces a classifier or model. Thus, PL is more straightforward and easier to implement.

Active machine learning (AML) is a popular research area in ML [5, 8, 13]. It allows selection of the most informative instances in training dataset of the domain for manual labeling. AML aims to produce a highly accurate classifier using as few labeled instances as possible, thereby minimizing the cost of obtaining labeled data [5]. With this assumption, AML is a specialized version of SSL, and they aim to reduce manual labeling workload. We provide a comparison of AML and PL in **Table 1** [5, 13]. Although most research in AL has tended to focus on binary classification problems and achieve high classification accuracy, some studies have addressed multicategory classification [15]. In AML, the classifier is initially trained on a small set of instances (labeled pool L), and it chooses informative instances from an unlabeled pool U of data to request labels from the expert or oracle (e.g., a human annotator) that can upon request provide a label for any instance in this pool [16], and this is repeated till convergence. This way, only a small quantity of unlabeled instances needs to be labeled to get very good performance. **Figure 3** illustrates the AML cycle.

AML is an important technique in machine learning, because labeled data is often more difficult and expensive to obtain than unlabeled data [17]. For example, if one is classifying web pages into categories based on the content, labeled data would likely be collected by hand, while unlabeled data could be found from the Internet automatically. Multiple studies proposed several AML algorithms and applied them to many applications. They have shown that when AML is used, ML models require significantly less training data and can still perform well without loss of accuracy [4, 8].

Typically, AML approaches select a single unlabeled instance, which is the most informative at that iteration, and then retrain the classifier. The training process in this case is hard and time-consuming; further, repeated retraining is inefficient. Thus, a batch mode AML strategy [9, 10] that selects multiple instances each time is more appropriate under these circumstances.

Pool-based method is the most prominent technique used in AML, and most research work of AML is pool based in recent years as unlabeled data has become easier to collect [5]. Pool-based AML assumes that the model has access to the entire set of unlabeled data at selection time.

PL	AML
No control over training instances	Selects training instances from a pool of unlabeled data (queries)
Large number of required training instances	Relatively small
Examine the entire training data before inducing a classifier (batch process)	Learner sees one or a subset instances at a time (iterative process)
One classifier induced	Many
Simple stopping criteria	Complex

Table 1. Comparison of active machine learning and passive learning.
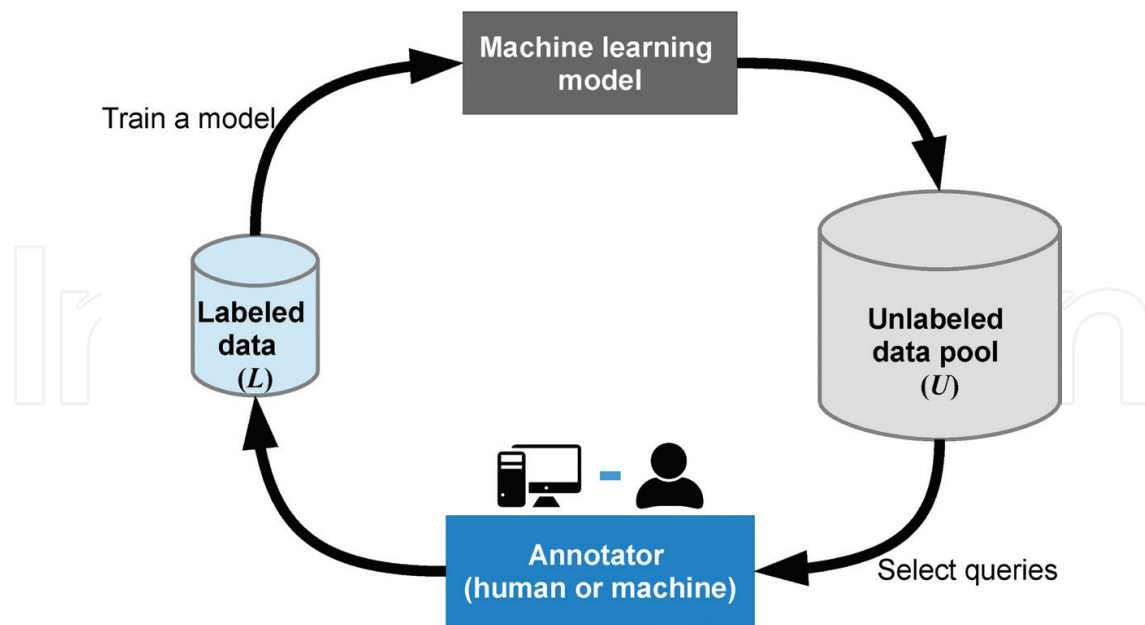
Figure 3. Pool-based AML cycle.

3. AML selection strategies

There are a number of AML query selection strategies, which have been presented by Settles [5]: (1) Uncertainty sampling is the simplest and the most commonly used strategy. Uncertainty sampling focuses on selecting the instance that the classifier is most uncertain about to label. This strategy can be divided into two categories: maximum entropy of the estimated label and minimum margin (distance of an instance to the decision boundary). (2) Expected error reduction, which aims to query instance that minimizes the expected error of the classifier. (3) Query by Committee (QBC) in which the most informative instance is the one that a committee of classifiers finds most disagreement. Bagging and boosting are used to generate committees of classifiers from the same dataset. They aim to combine a set of weak classifiers to create a single strong classifier. While bagging creates each base classifier independently, boosting allows these classifiers to influence each other during training process [18]. Boosting is an iterative process that initially assigns equal weight to each of the training samples; then the weights are modified based on the error rate of individual classifier.

4. Human AML

Actually human learns concepts in similar way of SSL, from a limited number of labeled data (e.g., parental labeling of objects) to a large amount of unlabeled data (e.g., observation of objects without naming in real life) [14]. In the ML scenario, it is easy to obtain the predictions of the classifier, and it is usually expensive to obtain the actual labels for instances.

In the real world, learners are not provided with labeled category information with every object they encounter (like in supervised category learning tasks), nor do they receive only unlabeled

information (like in unsupervised category learning tasks). People use labeled (with feedback) and unlabeled (no feedback) information when learning categories. In supervised learning, individuals learn the categories by correcting their performance based on the feedback they receive. The feedback people receive about categories may be either true or false. However, unsupervised learning category gives no feedback (information) about the category of an object. The individual learns from his/her experiences with different category objects, without receiving any feedback. Actually, humans categorize real-world categories most similarly to an SSL technique.

Gibson et al. [19] have used the equivalences between models found in human categorization and machine learning research to explain how the SSL techniques can be applied to human learning. In human AML, participants are usually first shown a small number of labeled instances, followed by a large set of unlabeled instances [20]. A set of experiments conducted by Gibson et al.[19] showed that SSL models are useful for explaining human behavior when they used both labeled and unlabeled data.

Unlike machines, a human learner gets tired as she/he answers questions, so finding out whether she/he knows a concept or not (i.e., getting the labels) is usually an expensive task. Therefore, using AML for human learning may help it become more efficient and effective and hence reduce the cost of teaching [21]. It might be costlier to teach an example to a human than it is to teach a computer program.

Although AML has been studied in different domains, such as video annotation and web page classification, its applications to human learning have been studied very little. There are a few empirical studies on the applications of AML to the human learning domain. The first such study was by [22], which showed that humans can use unlabeled data in addition to labeled data in categorization tasks. The authors in [23] proved by empirical evidence that human category learning is influenced by unlabeled data in a supervised categorization task, but they did not explain how individual can select these unlabeled examples.

There are a few studies on the applications of AML to the human category learning domain. Castro et al. [12] investigated what they refer to as "human active learning." They tried to answer a research question "Can machine learning be used to enhance human learning?" in the context of human category learning. We consider [16] as the most interesting study in human AML field. Castro et al. showed that humans learn faster with greater performance when they can actively select the instances from a pool of unlabeled data instead of random sampling and their performance is nearly optimal. However, they did not address how humans choose the next best instance. Moreover, they conducted their experiments by humans in a simple binary classification task, not in a real-life situation. Participants were presented with artificial novel 3D shapes (stimuli) that varied along a single, continuous dimension (spiky to smooth) and were given feedback as to which category the stimulus belonged to (see **Figure 4**). The task for each participant was to find out the precise egg shape (category boundary) for which eggs that were any spikier would hatch into snakes, while eggs that were any smoother would hatch into birds. Authors compared the performance of three distinct conditions for each participant. In the active learning condition, the participant could choose specific observations to test their beliefs based on her previous queries and their noisy labels, whereas, in
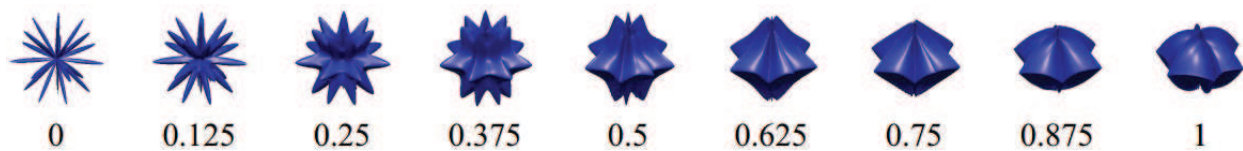
Figure 4. Example stimuli used in the experiment, with corresponding values.

the PL condition, the sequence of data was generated randomly by the experiment. They also included a "Machine-Yoked" condition where participants saw sequences of observations created by active learners but did not have control over the sequence.

The closest research to investigation by Castro et al. is [24] and [25] using a slightly modified procedure. They have successfully shown that learners benefit from the selection in category learning. Gureckis and Markant [25] concluded that AML can be superior because it allows humans to use their prior experience and current hypotheses to select the most helpful instances (e.g., asking a question about something that is especially confusing).

The work [24] by the same authors examined the interaction of self-directed information selection and category learning. Self-directed learning in humans can be inspired by "active learning" research in the ML literature [25, 26]. In this study, participants learned about two categories of "antennas" that varied along two dimensions (circles that differed in size and the orientation of a central line segment) and received one of the two television stations (CH1 or CH2). They compared active learning (or self-directed learning) condition, in which participants designed stimuli to learn about, with passive condition in which instances were generated from predefined distributions. Their results showed that for simple one-dimensional rules, active learners acquired the correct category rule faster than passive learners. Also, the AML advantage only held for the less complex, rule-based category.

Sim et al. [27] showed that school-age children learn more effectively when they are allowed to make decisions about what information they wish to gather than others who could only observe samples that were randomly generated for them. This results lead to the conclusion that children are capable of learning from the data they generate by themselves even at an early age. This result also suggests that the children's information gathering was informed by uncertainty and previous feedback, leading them to sample items that were near the true category boundary. This result was successfully replicated in [24]. Adams and Kachergis proved the effectiveness of AML for preschoolers, and they use an informative sampling strategy in an active category learning task. The authors suggest that children's performance in the AML task is related to their early math and preliteracy skills.

Kachergis et al. [28] in their paper investigated whether AML is better than PL in a cross-situational word learning context. They also investigated the strategies and found that most learners use immediate repetition to disambiguate pairings.

Researches in computer science on computationally efficient AML have inspired new theoretical approaches to inquiry behavior in humans.

5. Challenges of using AML for human categorization

5.1. High dimensionality

Most of the empirical studies, which addressed human HAML, focus on learning two-class problems on a one-dimensional input space [15, 29], but there are obstacles to generalize the model to multiclass classification problems [19]. The same is true for a single-modality (visual) object recognition. It is important to investigate the generalization of human AML to multi-dimensional objects (stimuli), such as auditory stimuli, high-dimensional stimuli, real-world stimuli, other demographic groups, etc.

5.2. Sensitivity to noise

Experimental studies on the human AML show that it is sensitive to noise and humans are not as good as machines in selecting queries from an unlabeled dataset of artificial visual stimuli [21]. Thus, human AML performance declines with higher noise levels. Humans perform relatively well in at least some noise settings, suggesting that they took the experiment seriously [30].

5.3. Distribution and ordering of unlabeled data

Zhu et al. [22] showed that humans are sensitive to the distributional structure of the subsequent unlabeled experience. Gibson et al. [19] investigated the effects of the distributions of unlabeled instances (stimuli) to human learner in two experiments, and they also investigated the effect of the order of the unlabeled items that participants encountered in an experiment. They concluded that human categorization is sensitive to both the distribution and ordering of unlabeled instances [19].

5.4. Small number of participants and objects

The small number of participants limits the generalization of the findings to other humans. In many studies that investigated human AML, small group of people participated in the experiments. In addition, a small number of objects used in the investigations lead to a similar limitation, because a limited number of teaching and test instances reduce the reliability of the results.

5.5. Category or concept structure

People are sensitive to the value of both labeled and unlabeled stimuli, and this depends on the structure of the concept being learned [24]. Markant and Gureckis [31] showed that the effectiveness of AML might interact with the particular structure of the target categories. Two types of category structures were used in that study: rule-based (RB), in which the decision rule is defined as a criterion along a single dimension, and information integration, in which the decision rule is a function of at least two dimensions.

5.6. Prediction of the next question

In human learning, people often learn by asking rich and interesting questions, which more directly target the concepts in a learning task. For example, a child might ask "Do all dogs have long tails?" or "What is the difference between cats and dogs?" [32]. The main challenge for AML method is to predict which question a human will ask from the given context. A number of recent studies have discussed this challenge [33–35]. Rothe et al. in [33] proposed a model that predicts what questions human learners will ask and can creatively generate novel questions that did not exist in the training data. Their work in [34] showed that human can accurately evaluate question quality by using the Bayesian ideal observers. In the most recent review [35], authors highlight and discuss nine challenges about the psychology of human inquiry.

6. Human AML for autism students

One goal of recent researches is to incorporate ML models with behavioral models to teach students and to investigate whether they can benefit from ML techniques to learn better.

Radwan et al. [21] were the first to attempt to use AML for teaching students with autism spectrum disorders (ASD). Students with ASD cannot learn in the same way as most people, and they need special treatment to learn a concept or an object. One of the difficulties faced by people with ASD is the recognition of categories. Radwan et al. [21] proposed a novel batch-mode pool-based AML framework for teaching students with ASD and compare the effectiveness of PL vs. AML on teaching object recognition for those students. AML approach presented to the student the most informative teaching set of objects based on the uncertainty sampling strategy. In this framework, a student plays the role of the classifier and does not have a probabilistic model. So, the uncertainty is computed in the context of the child's responses to measure informativeness for all objects. If an object's uncertainty is high, it implies that the student does not have sufficient knowledge to classify the object, and then adding this object into the training set can improve the child's recognition ability.

For this purpose, a web- and touch-based application was developed and presented on a tablet PC. Objects from everyday lives of children were grouped based on their categories and four difficulty levels L1–L4; see **Figure 5**. Picture stimuli of target objects were colored images, and they were collected using image search engines, in particular Google and Bing. The teaching procedure was based on applied behavioral analysis (ABA) principles. Five students with mild to moderate levels of ASD participated in the experiment. An alternating treatment design of single subject research methods was used to compare the effects of AML and PL.

The results indicate that AML was more effective than PL for four out of the five students. Consequently, students can learn faster and are able to reach a learning criterion with fewer teaching trials [21]. AML approach was generally more effective in terms of accuracy. The statistical results demonstrated that there was a statistically significant difference in accuracy level between the means of PL and AML. The AML approach and procedures provide two
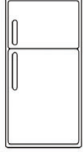
Figure 5. A sample of images from the dataset used for teaching and test. Columns L1–L4 show different difficulty levels.

features that helped to reduce repetition in learning environment: (1) minimizing the number of teaching trials required for training and (2) determining mastery criterion for levels. When a participant reached mastery criterion, the application no longer assesses this level in the following phases.

7. Conclusion

The applications of AML in human categorization have become increasingly common in recent years. Humans and machines seem to benefit from AML in similar ways [24, 30, 36]. In AML setting, a learning machine is able to query an oracle in order to obtain the most informative instances that are expected to improve performance [37]. However, humans ask far richer and more sophisticated questions. In this review, we present and discuss benefits and challenges of using AML for human categorization and concept learning. More research is needed to address several limitations of human AL.

Author details

Akram M. Radwan

Address all correspondence to: akr.radwan@gmail.com

Department of Information Technology, University College of Applied Sciences, Gaza, Palestine

References

[1] Sinha P, Sinha P. Comparative Study of Chronic Kidney Disease Prediction using KNN and SVM [Online]. Available from: paper/Comparative-Study-of-Chronic-Kidney-Disease-using-Sinha-Sinha/3ec05afd1eb4bb4d5ec17a9e0b3d09f5cbc30304

[2] O'Neill J. An evaluation of selection strategies for active learning with regression [dissertations thesis]. 2015

[3] Bird S, Klein E, Loper E. Natural Language Processing with Python. 1st ed. O'Reilly Media, Inc; 2009

[4] Nore PW. Pollution detection in a low-cost electronic nose, a machine learning approach [thesis]. 2016

[5] Settles B. Active Learning Literature Survey. Vol. 52, No. 55-66. Madison: University of Wisconsin; 2010. p. 11

[6] Kremer J, Pedersen KS, Igel C. Active learning with support vector machines. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2014;**4**(4):313-326

[7] Wuest T, Weimer D, Irgens C, Thoben KD. Machine learning in manufacturing: Advantages, challenges, and applications. Production & Manufacturing Research. **4**(1): 23-45. DOI:10.1080/21693277.2016.1192517

[8] Tong S. Active Learning: Theory and Applications. Stanford University; 2001

[9] Chen Y, Mani S, Xu H. Applying active learning to assertion classification of concepts in clinical text. Journal of Biomedical Informatics. 2012;**45**(2):265-272

[10] Grabinger S, Dunlap JC, Duffield JA. Rich environments for active learning in action: Problem-based learning. Alternatives Journal. 1997;**5**(2):5-17

[11] Deng C, Liu X, Li C, Tao D. Active multi-kernel domain adaptation for hyperspectral image classification. Pattern Recognition. 2018;**77**(C):306-315

[12] Castro RM, Kalish C, Nowak R, Qian R, Rogers T, Zhu X. Human active learning. In: Advances in Neural Information Processing Systems (NIPS), Vancouver, Canada. 2008. pp. 241-248

[13] Davy M. A review of active learning and co-training in text classification. In: Computer Science Technical Report TCD-CS-2005-64. Dublin: Trinity College; 2005. pp. 170-179

[14] Zhu X, Goldberg AB. Introduction to Semi-Supervised Learning, Synthesis Lectures on Artificial Intelligence and Machine Learning. San Refael: Morgan & Claypool Publishers; 2009;**3**(1):1-130

[15] Schein AI, Ungar LH. Active learning for logistic regression: an evaluation. Machine Learning. 2007;**68**(3):235-265

[16] Tong S, Koller D. Support vector machine active learning with applications to text classification. Journal of Machine Learning Research. 2002;**2**:45-66

[17] Chapelle O, Scholkopf B, Zien EA. Semi-supervised learning (Chapelle O. et al., eds. 2006) [Book reviews]. IEEE Transactions on Neural Networks. 2009;**20**(3):542-542

[18] Olsson F. A literature survey of active machine learning in the context of natural language processing [Online]. 2009. Available from: http://eprints.sics.se/3600/

[19] Gibson BR, Rogers TT, Zhu X. Human semi-supervised learning. Topics in Cognitive Science. 2013;**5**(1):132-172

[20] Zhu X, Gibson BR, Jerryzhu KJ, Edu DW, Rogers TT, Harrison J, et al. Cognitive Models of Test-Item Effects in Human Category Learning. In: Proceedings of the 27th International Conference on International Conference on Machine Learning, June 21-24. 2010. pp. 1247-1254

[21] Radwan AM, Birkan B, Hania F, Cataltepe Z. Active machine learning framework for teaching object recognition skills to children with autism. International Journal of Developmental Disabilities. 2017;**63**:158-169

[22] Zhu X, Rogers T, Qian R, Kalish C. Humans perform semi-supervised classification too. In: Proceedings of the National Conference on Artificial Intelligence. AAAI Press. 2007;**22**:864-869

[23] Guo Y, Schuurmans D. Discriminative batch mode active learning. In: Advances in Neural Information Processing Systems (NIPS), Cambridge, MA: MIT Press; 2008;**20**:593-600

[24] Markant DB, Gureckis TM. Is It Better to Select or to Receive? Learning via Active and Passive Hypothesis Testing. Journal of Experimental Psychology: General. 2014;**143**(1):94-122

[25] Gureckis TM, Markant DB. Self-directed learning: A cognitive and computational perspective. Perspectives on Psychological Science. 2012;**7**(5):464-481

[26] Prince M. Does active learning work? A review of the research. Journal of Engineering Education. 2004;**93**(3):223-231

[27] Sim ZL, Tanner MM, Alpert NY, Xu F. Children learn better when they select their own data, CogSci [Online]. 2015. Available from: paper/Children-Learn-Better-When-They-Select-Their-Own-Sim-Tanner/c785ae99b502e4c068ae128faeb9bb7d27f6dbaf

[28] Kachergis G, Yu C, Shiffrin RM. Developing semantic knowledge through crosssituational word learning. In: Proceedings of the 36th Annual Conference of the Cognitive Science Society, Quebec City, Canada, July 23-26, 2014. Austin, TX; 2014

[29] Yang Y, Ma Z, Nie F, Chang X, Hauptmann AG. Multi-class active learning by uncertainty sampling with diversity maximization. International Journal of Computer Vision. 2014;**113**(2):113-127

[30] Castro RM, Kalish C, Nowak R, Qian R, Rogers T, Zhu X. Human active learning. Proc. of the 22nd annual conference on neural information processing systems. Vancouver, British Columbia, Canada, 2008, pp. 241-248

[31] Markant D, Gureckis TM. Category learning through active sampling. In: Ohlsson S, Catrambone R, editors. Proceedings of the 32th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society; pp. 248-253

[32] Siegler RS. Mechanisms of cognitive development. Annual Review of Psychology. 1989; **40**:353-379

[33] Rothe A, Lake BM, Gureckis T. Question asking as program generation. In: Advances in Neural Information Processing Systems 30 (NIPS), 2017. pp. 1046-1055

[34] Rothe A, Lake BM, Gureckis T. Do people ask good questions? PsyArXiv. 2018. DOI: 10.31234/osf.io/eanku

[35] Coenen A, Nelson JD, Gureckis TM. Asking the right questions about the psychology of human inquiry: Nine open challenges. Psychonomic Bulletin & Review. 2018. pp. 1-41. DOI: 10.3758/s13423-018-1470-5

[36] Rothe A, Lake BM, Gureckis TM. Asking and evaluating natural language questions. In: Papafragou A, Grodner D, Mirman D, Trueswell J, editors. Proceedings of the 38th Annual Conference of the Cognitive Science Society, Philadelphia, US, August 10-13, 2016. Austin, TX; 2016

[37] Hoi SC, Jin R, Zhu J, Lyu MR. Batch mode active learning and its application to medical image classification. In: Proceedings of the 23rd International Conference on Machine Learning (ICML2006), Pittsburgh, PA, US, June 25-29, 2006. pp. 417-424