COMP7703 Machine Learning
Article Review 1

Thanat Chokwijitkul 44522328

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78-87.

The article, "A few useful things to know about machine learning" by Pedro Domingos discusses various lessons learned by machine learning researchers and practitioners. It provides the information on essential issues to focus on and pitfalls to avoid, along with useful recommendations.

Throughout the article, Domingos only focused on classification for the purpose of illustration. The author outlined various critical machine learning related issues which impact the development of successful classification applications. Most of these issues primarily concentrate on essential key components, including data and its features, and learning algorithms (classifiers).

Domingos also introduced two of the leading problems, including overfitting and the curse of dimensionality, as common pitfalls to avoid. Another unobvious pitfall is the use of theoretical guarantees. Those guarantees should only be used as a driving force for learning algorithm design, not a criterion for decision support.

Essentially, the author emphasised that the ultimate goal of machine learning is to generalise beyond the given examples in the training data because exact examples are unlikely to appear at test time.

It is evident that Domingos has considerable expertise and deep experience in this subfield of artificial intelligence. Apart from summarising the general ideas of machine learning in fairly approachable terms, Domingos also managed to divulge the importance of "folk knowledge" or common sense in the field. The author's primary intention was to identify various issues which have been commonly overlooked or misunderstood, along with providing recommendations to assist researchers in developing more efficient machine learning systems.

However, there are still some doubts about the practicality of some concepts presented in the article. The case of Domingos' criteria for classifier selection (representation, evaluation and optimisation) is an excellent example for this. Theoretically, those criteria are consistent and well-constructed. Nevertheless, in practice, the set of possible models for the real learning algorithms is extremely massive. One of the challenges is how to find a good one without trying all the existing algorithms. Seeking help from machine learning experts is a useful method to make this process less

time-consuming, or conducting intensive research on the subject with more time spent. However, these approaches may not result in finding the best algorithm, but at least a moderately good one.

Apart from that, everything else is very well-justified, particularly on the topic of pitfalls. Domingos' inclusion on the two common problems brings the useful technical insight for machine learning researchers and practitioners on why many learning algorithms fail while other components are well-constructed.

Additionally, the author's opinion on theoretical guarantees is not overstated. It is indeed an excellent argument regarding the real-world impacts of those theoretical guarantees since many tend to be sensitive with a fancy mathematical proof behind an algorithm, which leads to the thought that it must be better than all others.

Overall, Domingos' work provided the comprehensively insightful information on several issues confronted by machine learning researchers and practitioners. As discussed above, the article covers important issues to focus on along with common risks to avoid. Nevertheless, apart from the content presented in this article, there definitely will be many more lessons to learn in the future since machine learning is still far from its end, and what exists now is just the humble beginning.