

# Emperical Analysis of General utility problem in Machine Learning

Lawrence B. Holder  
holder@cse.uta.edu

Department of Computer Science Engineering  
University of Texas at Arlington  
Box 19015, Arlington, TX 76019-0015

## 1 Pros and Cons of the paper

### 1.1 Pros

- This paper is a good mix of observation's over a range of machine learning paradigms and not a few thereby helping us to generalize the utility problem. Also it helps us to verify if the same trend/problem exists in all model's or not with practical application of the method.
- The paper helps show that across all methods discussed i.e. Neural network method, splitting method, set-covering method etc. they follow a some what general trend i.e. trend of the utility problem and helps us generalize the utility problem
- This paper also gives a basic overview of each of the method's involved, thereby clearing the readers vision and also helping him understand each of the methods discussed. Also, the paper has two table's that lists the entries for several of the inductive learning and Planner speedup learner and is extremely useful
- When applying some model and drawing graph, proper value of the axis and the reason for such a choice is clearly given for each of the graphs. The important conclusion from the graph is also mentioned. Thereby, no stone is left unturned to explain the concept involved to the reader.
- The formal model helps realize and complete the understanding of the general utility problem and also it formalizes the approach to the problem with the introduction of the Bayes optimal classification error and variance. It also helps to identify some of the obvious solution to the problem e.g. Stopping the learning process when peak is reached. The identification and the max peak possible to be obtained is further helped by the table shown for various method on different domains.
- The concept of bias and variance has been linked well to the no. of leaves  $L$  in the decision tree and was fundamental in understanding and representing the formal model of the performance response for splitting methods.
- In speed up learner topic, the relation of the no. of macro-operators or control rules minimizing the time taken by the problem solver to solve problems and inverse CPU

time for planner is also given apart from that for inductive learner. Thereby further helping us in generalizing the utility problem.

## 1.2 Cons

- The paper has not given quite a lot of method's when it comes to solving the problem of the gradual decline of performance as the amount of learned knowledge increases but has only suggested stopping at an early point or at the point of giving the peak as a probable solution to the problem. Other methods may also be there
- The paper although has listed the MBAC model. But it has not given solution to and or explained the topic as clearly as of the other topics such as of Inductive learners.
- Regarding the peak performance on various domains, the reason for huge differences sometimes in result has not been discussed even though there is a significant difference in the highest percentage and lowest percent of peak performance amongst the domain and methods discussed. Also, some domain are exceeding the maximum 100% limit and no explanation of the result is given.
- This paper gave a lot of information regarding the amount of learned knowledge to that of the performance. But any effect of the size of the data and the computing power of the device on the value of the best performance percentage was not given.
- Many of the models example. neural network model have not one but many local maximums when plotting the performance response. If a proper graph is plotted then its easy to find the global maximum but this is not the case when actually implementing the model. No such method has been explained for finding the global maximum in case there are many local maximum.
- Not all systems suffer from the utility problem; Static machine learning systems don't generally suffer from utility problem. The utility problem only arises in dynamic learning systems. Moreover, it primarily arises in systems that have a secondary goal of speeding up their processing on some task other than the learning task itself. This has not been discussed in the paper.

## 2 Technical suggestions for improving some component of the technique or experimental analysis

- Here we see that many of the methods such as neural network method in case of accuracy vs number of cycles and also speed up learning on the no. of macro-ops vs inverse CPU time has many local maximas. So, we have to make sure that we check for the performance of the model well beyond the point of maximum, so as to make sure that the global maximum is reached. Other wise, it may happen that the performance percentage obtained is for some local maximum and not for the global maximum.
- As we see that in the splitting method, when using different tree traversal methods, undoubtedly, breadth first has the highest performance. But when we want a trade off between the amount of data learnt and the performance for example no. of split and the performance. Other methods can become more effective over larger range of data learnt for example larger no. of splits.

### 3 Further Directions For Research

- A way forward from this would be to find the effect the size of the data and the computing power of the device on the value of the best performance percentage as this would also help to find the best combination while training.
- Also, a research on refinement of models of the general utility problem will help provide a general framework for controlling and comparing different learning paradigms and therefore will help in advancement of the model and predicting the needed information without a decrease in performance.
- from this as we see that for the MBAC model, it uses empirical model of performance response. Therefore, the MBAC approach or any approach to controlling and estimating the performance of a learning method would benefit from a formal model of the performance response that depends on properties of the current learning task and a research on such approaches and models would help in progress.
- Another direction could be finding the best method out of the ones used in the paper example ID3, PLS1 etc for each domain.