Performance Evaluation in Machine Learning: The Good, The Bad, The Ugly and The Way Forward

**ABSTRACT**

This paper gives an overview of some ways in which our understanding of performance evaluation measures for machine-learned classifiers has improved over the last twenty years. I also highlight a range of areas where this understanding is still lacking, leading to ill-advised practices in classifier evaluation. This suggests that in order to make further progress we need to develop a proper measurement theory of machine learning. I then demonstrate by example what such a measurement theory might look like and what kinds of new results it would entail. Finally, I argue that key properties such as classification ability and data set difficulty are unlikely to be directly observable, suggesting the need for latent-variable models and causal inference.

**EXISTING SYSTEM**

Measurement is of evident importance in machine learning in at least two ways: first is the features used by machine learning models are themselves measurements, which is a perspective explored .second performance evaluation of learned models, which was the topic of this paper. The development of a measurement theory for machine learning is lagging behind the remarkable achievements of machine learning technology itself, a situation that is in urgent need of rectification if we want that technology to be accepted and trusted by users.

**PROPOSED SYSTEM**

This suggests that in order to make further progress we need to develop a proper measurement theory of machine learning. I then demonstrate by example what such a measurement theory might look like and what kinds of new results it would entail. I will start by reviewing and critiquing current practice in performance evaluation of machine learning algorithms. I will highlight some good things, some not so good things, and some things to be avoided. This is intended to demonstrate, by example, the need for a more careful treatment of performance evaluation and the development of a specific measurement framework for machine learning, but should not be taken as complete in any sense. It is hoped that this paper will provide an impetus in that direction.

# Hardware Requirements:

# Processor - Pentium –IV

* Speed - 1.1 GHz
* Ram - 256 MB
* Hard Disk - 20 GB
* Key Board - Standard Windows Keyboard
* Mouse - Two or Three Button Mouse
* Monitor - SVGA

**Software Requirements:**

* Operating System - Windows XP
* Coding Language - Java