2020

36th IARIW General Conference

Paper Prepared for the 36th IARIW General Conference, Oslo, Norway, August 24-28, 2020

Metrics for Measuring the Performance of Machine Learning Prediction Models: An Application to the Housing Market

Miriam Steurer

Robert Hill

There has been a recent surge of interest in machine learning (ML) prediction methods. While cross-validation (CV) is typically used for model selection, metrics are still required to evaluate predictive performance at each stage of the CV exercise. The existing literature is interdisciplinary, making it hard to compare the available metrics. We collect the most commonly used metrics, classify them by type, and then evaluate them with respect to two novel symmetry conditions. While none of these metrics satisfy both conditions, we propose a number of new metrics that do. Our findings are illustrated with an application in which five ML methods are used to predict apartment prices. We show that the most popular metrics in the automated valuation literature generate misleading results. A different picture emerges when the full set of metrics is considered, and especially when we focus on four key metrics with the best symmetry properties.