

A framework for monitoring classifiers performance: when and why failure occurs

Nitesh V. Chawla

University of Notre Dame, USA

Abstract. Classifier error is the product of model bias and data variance. While understanding the bias involved when selecting a given learning algorithm, it is similarly important to understand the variability in data over time, since even the One True Model might perform poorly when training and evaluation samples diverge. Thus, the ability to identify distributional divergence is critical towards pinpointing when fracture points in classifier performance will occur. Contemporary evaluation methods do not take the impact of distribution shifts on the quality of classifiers predictions. In this talk, I present a comprehensive framework to proactively detect breakpoints in classifiers predictions and shifts in data distributions through a series of statistical tests. I outline and utilize three scenarios under which data changes: sample selection bias, covariate shift, and shifting class priors.