

Machine Learning for Imbalanced Data and Applications to Severe Weather Prediction

Theodore B. Trafalis

Professor, School of Industrial and Systems Engineering, The University of Oklahoma, Norman,
OK 73019, USA

Email: ttrafalis@ou.edu

Abstract: Imbalanced learning has taken a lot of attention in recent years in machine learning. Imbalanced learning is the art of extracting knowledge from a data set where the number of instances in one class called the majority class is significantly larger than the number of instances in the other class known as the minority class. The minority class consists of rare cases that are more important from the learning perspective. For example in the case of severe weather, tornados refer to the minority class. The main concern in imbalanced learning is that most classifiers are biased towards the majority class, and they fail to classify the instances in the minority class correctly. In this tutorial we discuss several machine learning approaches for imbalanced data and applications to extreme weather events such as tornados.