

Órdenes estocásticos aplicados a problemas de aprendizaje de máquina con valores de pérdida no numéricos.

Inés Couso, Luciano Sánchez.

Universidad de Oviedo.

Abstract: Se estudian distintos métodos de evaluación para seleccionar la decisión (modelo) óptima en problemas de regresión y clasificación. Se proponen alternativas al clásico criterio de la minimización de la pérdida esperada. En nuestro contexto, identificamos cada modelo de regresión o clasificación con una variable aleatoria que asocia, a cada individuo de la población, un valor de pérdida (distancia cuadrática o distancia absoluta en problemas de regresión, pérdida binaria en problemas de clasificación, etc.) Sobre la base de esta identificación, cada orden estocástico da lugar a un posible criterio de comparación entre los diferentes modelos, de los cuales, el criterio basado en la minimización de la pérdida media es solo un caso particular. Órdenes estocásticos alternativos al de la comparación de las esperanzas, tales como la dominancia estocástica de primer orden o el orden basado en el criterio de Condorcet no requieren que la función de pérdida sea numérica. Pero incluso en el caso de que esta última sea numérica, pueden dar lugar a modelos más robustos que aquellos que optimizan la pérdida esperada. En este trabajo se presentan e ilustran algunos de esos criterios alternativos.

Palabras Clave: Órdenes estocásticos; Regresión; Clasificación; Función de pérdida ordinal.