

Métodos de clasificación supervisada en series temporales

Sofia Ruiz Suarez¹, Vianey Leos Barajas^{2,3}, Juan Manuel Morales¹

¹ INIBIOMA-CONICET, Bariloche, Argentina.

² Iowa State University, USA

³ Bielefeld University, Germany

Palabras Claves: *dependencia temporal, machine learning, simulaciones.*

Una serie de tiempo se puede definir como una secuencia de medidas a lo largo del tiempo que describen el comportamiento de cierto sistema. El mismo puede cambiar a lo largo del tiempo debido a cambios tanto internos como externos, permitiendo así distinguir diferentes estados o condiciones del sistema. El problema de clasificación en este caso tiene como objetivo tanto reconocer los instantes de cambio de estado en el sistema como las categorías de los mismos. Existe una variedad de campos en donde este problema es aplicable, por ejemplo en medicina para reconocer distintos ritmos o frecuencias cardíacas de un paciente según un electrocardiograma; en computación para el problema del reconocimiento de voz e incluso en ecología para el reconocimiento de comportamientos en animales a partir de acelerómetros y GPS.

Diferentes enfoques y metodologías pueden ser utilizados con este propósito, y dependerá de los datos disponibles y la naturaleza de los mismos su verdadera utilidad. Metodologías del tipo *Machine Learning*, tales como *Random Forest* o *Support Vector Machines* han sido utilizadas con éxito para resolver problemas de clasificación. Sin embargo, estas metodologías tal cual se presentan no consideran la temporalidad de los datos, por lo que es necesario agregar esta condición dentro del análisis. Los modelos *hidden Markov model* son una alternativa para este problema. Estos modelos de serie de tiempo estocástico incluyen la configuración temporal dentro de su estructura, lo que los hace una alternativa potente para resolver este problema.

Mediante simulaciones en **R** comparamos estos dos enfoques y exponemos las fortalezas y limitaciones de ambos. Estudiamos las consecuencias de aplicar técnicas suponiendo independencia temporal y analizamos los escenarios en donde es posible aplicar estas técnicas para la clasificación de series temporales.