



## **APLICACIÓN DE TÉCNICAS DE MINERÍA DE DATOS, PARA LA GENERACIÓN DE PRONÓSTICOS DE DISPONIBILIDAD DE AGUA EN LA CUENCA BAJA DEL RÍO SAN LORENZO**

ANA LAURA HERRERA PRADO<sup>1</sup>, DIEGO URIBE AGUNDIS<sup>2</sup> y ARTURO RUÍZ LUNA<sup>3</sup>

1 Instituto Tecnológico de Mazatlán, 2 Instituto Tecnológico de la Laguna, 3 Centro de Investigación en Alimentación y Desarrollo, A.C.. anlauherrera@gmail.com

La Minería de Datos tiene un papel muy importante como tecnología de apoyo para explorar, analizar, comprender y aplicar el conocimiento adquirido de grandes volúmenes de datos, así como para identificar tendencias y comportamientos en la información, que faciliten una mejor comprensión de los fenómenos que nos rodean y ayudan en la toma de decisiones. En este estudio, se verificó la viabilidad del uso de diversas técnicas de minería de datos, como una herramienta computacional para predecir la disponibilidad de agua para fines agrícolas en la cuenca baja del río San Lorenzo. Este trabajo se basó en una revisión bibliográfica sobre estos temas y el análisis de repositorios disponibles en diversas instituciones mexicanas. Posteriormente, se definió, seleccionó y generó un dominio de datos. El cual se pre-procesó y transformó para entender el significado de los atributos, detectar errores de integración y lograr un dataset adecuado para construir modelos de series de tiempo. Las técnicas predictivas evaluadas fueron: Máquina de Soporte Vectorial, Regresión Lineal y Perceptrón Multicapas. El análisis y comparación de los patrones en series temporales arrojó como resultados: a) el escurrimiento del río San Lorenzo muestra dos comportamientos en periodos de tiempo distintos debido a la construcción y operación de la presa "José López Portillo" y b) las lluvias y evaporaciones registradas tienen un comportamiento estacional característico, reflejando los fenómenos meteorológicos esporádicos como huracanes o tormentas que ocurren en la región. Se obtuvieron correlaciones significativas al contrastar los resultados de las predicciones de las técnicas estudiadas, con los datos observados. Los coeficientes de correlación más altos se obtuvieron con la técnica Regresión Lineal. En conclusión, es viable el uso de técnicas de predicción para aplicarse en la generación de pronósticos de disponibilidad de agua hasta por un periodo de tres años, en la cuenca baja del río San Lorenzo.