



Camilo Andrade Murcia
Código 614122004
3 de octubre de 2020

Propuesta de proyecto final

Forecasting para ventas diarias de un restaurante usando Machine Learning

Los datos de ventas de un punto de venta suelen organizarse y analizarse desde el punto de vista de *series* de tiempo, área en el que hay varios modelos estadísticos que describen y permiten predecir el comportamiento de la serie de tiempo, como modelos auto regresivos AR, modelos de media móvil MV, sus combinaciones ARMA, ARIMA, SARIMA (cuando la serie de tiempo tiene un comportamiento periódico o de estaciones), ARIMAX cuando se incorpora variables externas a los datos de la serie de tiempo. Para ciertos casos estos modelos no producen buenas predicciones así que un tratamiento del problema usando Machine Learning es una alternativa para mejorar estas predicciones.

En Machine Learning para el análisis de series de tiempo, el programa diseñado aprende y predice el comportamiento de los datos basado en los datos pasados, para ello se usan técnicas de aprendizaje supervisado, árboles de decisión, redes neuronales artificiales y máquinas de vectores.

Objetivo

El objetivo de este proyecto es, para una base de datos de ventas diarias de un restaurante y su serie de tiempo asociada, crear un algoritmo con métodos de Machine Learning que permita predecir el futuro de la serie, y comparar los resultados con un modelo tradicional.

Acercamiento al problema

Los datos, además del valor diario de la venta del restaurante, tiene inmerso la información de la fecha, es decir, el año, el mes, el día de la semana, el # de día del mes, si es festivo o nó, que festividad es. La aproximación al problema que planteo es dadas esas clasificaciones de los datos es construir un predictor para la venta diaria de acuerdo a esas variables, haciéndolo con una parte de la serie de tiempo, de manera que primero aprenda el comportamiento de los datos por las frecuencias o periodicidades primarias, día de semana y mes, luego incorporar las excepciones de los festivos ya que son menor la cantidad de datos, y luego re alimentar el algoritmo para modelar la tendencia, debido a que la serie no es estacionaria.

Referencias

- [1] Zhang, G. P. (Ed.). (2004). Neural networks in business forecasting. IGI global.
- [2] Bontempi, G., Taieb, S. B., & Le Borgne, Y. A. (2012, July). Machine learning strategies for time series forecasting. In European business intelligence summer school (pp. 62-77). Springer, Berlin, Heidelberg.
- [3] Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.
- [4] Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1-2), 307-319.
- [5] Tay, F. E., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *omega*, 29(4), 309-317.