

# Predicción de $R(t)$ usando Movilidad

<sup>1</sup>INFOTEC Centro de Investigación e Innovación en Tecnologías de la Información y Comunicación, Circuito Tecnopolo Sur No 112, Fracc. Tecnopolo Pocitos II,

Aguascalientes 20313, México

<sup>2</sup> Centro de Investigación en Matemáticas, CIMAT-CONACYT, Jalisco S/N, Valenciana, Guanajuato, GTO, Mexico.

<sup>3</sup>CONACYT Consejo Nacional de Ciencia y Tecnología, Dirección de Cátedras, Insurgentes Sur 1582, Crédito Constructor, Ciudad de México 03940 México

## 1. Introducción

El **objetivo** es generar un modelo capaz de predecir la mediana del  $R(t)$ <sup>1</sup> utilizando como variable exógena los datos de movilidad.

El proceso de predicción se basa en un modelo autorregresivo de media móvil (ARMA), implementado en la librería statsmodels<sup>2</sup>, el cual usa como variable exógena la movilidad publicada por Google<sup>3</sup>. Cabe mencionar que los datos publicados por Google tiene un retraso de aproximadamente 5 días y por este motivo son actualizados mediante la movilidad estimada con datos de Twitter<sup>4</sup> para tener la información actualizada al día.

El proceso que se realizó fue el siguiente. El primer paso fue calcular el  $R(t)$  utilizando la información proporcionada por los datos abiertos de la Dirección General de Epidemiología. A la información publicada se le quitaron los últimos once días y se hizo el cálculo del  $R(t)$ . Una vez obtenido este, se procedió a obtener los valores  $p$  y  $q$  del modelo ARMA; estos parámetros fueron aquellos que se obtuvieron en una búsqueda donde el valor máximo de  $p$  es 4 y de  $q$  es 2. Se seleccionó los parámetros que minimizaban el Criterio Bayesiano de Información (BIC). Una vez obtenido los parámetros del modelo ARMA, se entrenó el modelo con los datos de  $R(t)$  y los correspondientes de movilidad. Finalmente, se utilizó el modelo ARMA para predecir los 14 días que se habían quitado, mas 7 días del desplazamiento en la movilidad y se incluyó en el estudio los intervalos de confianza de dicha predicción con una confianza del 95 %.

Es pertinente comentar que el modelo realizado utiliza para la predicción los valores previos del variable endógena, es decir,  $R(t-1), \dots, R(t-p)$ , los errores y los valores de movilidad actuales. Esto puede tener como consecuencia que un cambio drástico en la movilidad, que no fue observado en el entrenamiento, produzca un pronóstico no satisfactorio. Por ejemplo, un incremento en la movilidad drástico tendrá como consecuencia un incremento en el  $R(t)$  pronosticado, pero este incremento podría ser mayor que el medido, esto por la dinámica de la misma epidemia, es decir, otros factores que el modelo ARMA no está considerando.

<sup>1</sup> $R(t)$  se calcula usando <https://doi.org/10.1093/aje/kwt133> implementado por J. Andres Christen

<sup>2</sup><https://www.statsmodels.org/stable/index.html>

<sup>3</sup><https://www.google.com/covid19/mobility/>

<sup>4</sup><https://text-models.readthedocs.io/en/latest/place.html>

## 2. Predicción

En las siguiente figuras se muestra el  $R(t)$  calculado con los datos del 6 de agosto del 2020 así como la predicción del  $R(t)$  con sus valores de confianza al 95 %. La predicción llega hasta el día que fueron publicados los datos. En las figuras también se muestra la movilidad en gris claro y los valores de movilidad son apreciados en el lado derecho de cada figura.









