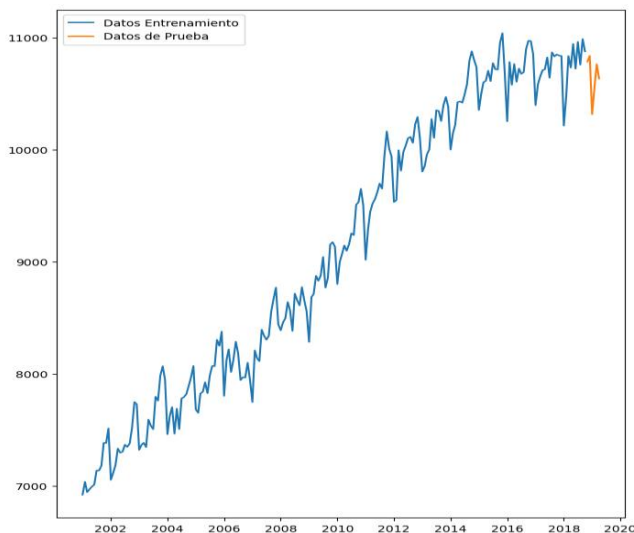


PRONOSTICO – PERSONAS OCUPADAS DE 13 CIUDADES PRINCIPALES DEL PAÍS / Raul Echeverry – Esteban Ordoñez

Los datos corresponden al número de personas ocupadas de 13 ciudades principales del país, que comprenden desde enero de 2001 hasta abril de 2019. El objetivo es pronosticar la cantidad de personas ocupadas en los meses de mayo a octubre de 2019, a través del uso de modelos ARIMA.

Protocolo de evaluación

Para lograr pronosticar los 6 meses siguientes a abril 2019, se realiza un protocolo de evaluación para entrenar y escoger el mejor modelo, esto implica usar los últimos seis meses, es decir desde nov 2018 a abril 2019.



Selección de Modelo

Se utiliza la función AUTO ARIMA para encontrar tanto el grado de AR, MA con o sin integración, dando como resultado: ARIMA(1,1,1):

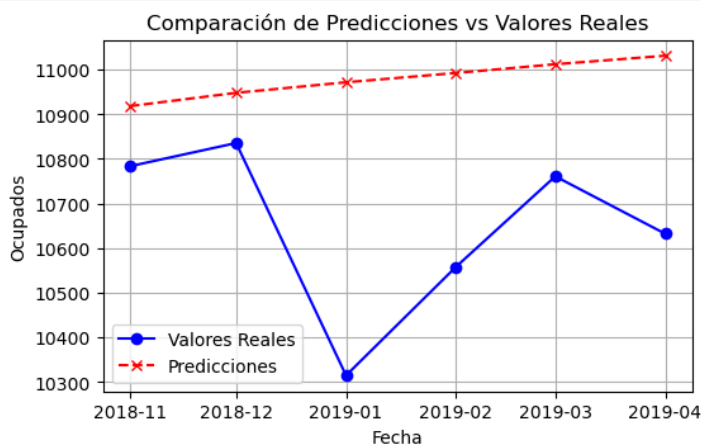
-AR1 (Autorregresivo): Significa que el modelo usa una observación retrasada (lag) para predecir el valor actual.

- I1 (Integrado): Significa que los datos han sido diferenciados una vez para hacerlos estacionarios, eliminando la tendencia en los datos y ayudando a estabilizar la media de la serie.

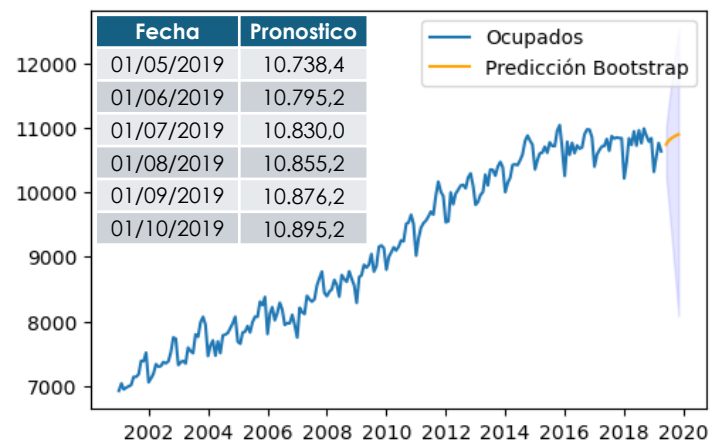
- MA1 (Media móvil): se usa media móvil basado en una observación retrasada del error del modelo (el residuo de la predicción anterior) para predecir el valor actual.

RMSE ARIMA (1,1,1): 381.7

Datos reales Vs Pronostico ARIMA(1,1,1)



Predicción final e intervalos de confianza (Usando Bootstrap)



Limitaciones: En el modelo desarrollado, se identificó una limitación importante relacionada con el incumplimiento del supuesto de normalidad de los residuos. Para mitigar esta limitación, se utilizaron intervalos de confianza generados mediante el método de bootstrap, lo cual permitió manejar la incertidumbre en las predicciones de manera más robusta.

Conclusión: En cuanto a la comparación con otros enfoques, los modelos Lineal de Holt y Lineal de Winters (Holt-Winters) lograron un RMSE de 93.5 en sus mejores configuraciones. Posteriormente, el modelo de polinomio de grado 5 mejoró este resultado, alcanzando un RMSE de 91.01. Sin embargo, el enfoque ARIMA implementado en este análisis, pese a los esfuerzos, presentó un RMSE significativamente mayor, de 381. Este resultado sugiere que, aunque el ARIMA es un modelo poderoso en muchas aplicaciones, en este caso específico no fue la opción más eficaz en comparación con los modelos anteriores.