

KNN TOTALES SEMANAL

Miguel Angel Villegas

2025-03-10

```
library(readxl)
library(forecast)

## Registered S3 method overwritten by 'quantmod':
##   method      from
## as.zoo.data.frame zoo

library(TSstudio)
library(tsfknn)
```

Introducción

Se descargan los valores de los productos ,luego se crea la serie temporal con una frecuencia mensual. Puesto que es un método no-parámtrico no existe información de AIC y de los residuales del modelo resultante. Por lo que el criterio a considerar para evaluar el modelo seran las metricas de exactitud, además el mejor modelo se determina en base al número de “k” vecinos en un vector de valores, que dependiendo el número de retrasos se puede maximizar pues el modelo tolera un número máximo de ambos combinados, en caso de sobre pasar el numero máximo de retrasos o vecinos próximos se despliega una advertencia la cual indica la imposibilidad de realizar el cálculo.

Se obtienen los datos

```
ruta_totales <- "/cloud/project/totales_transformados.xlsx"
excel_sheets(ruta_totales)

## [1] "Sheet 1"

totales_transformados <- as.data.frame(read_xlsx(path = ruta_totales, col_names = T,
                                                sheet = "Sheet 1"))
head(totales_transformados)

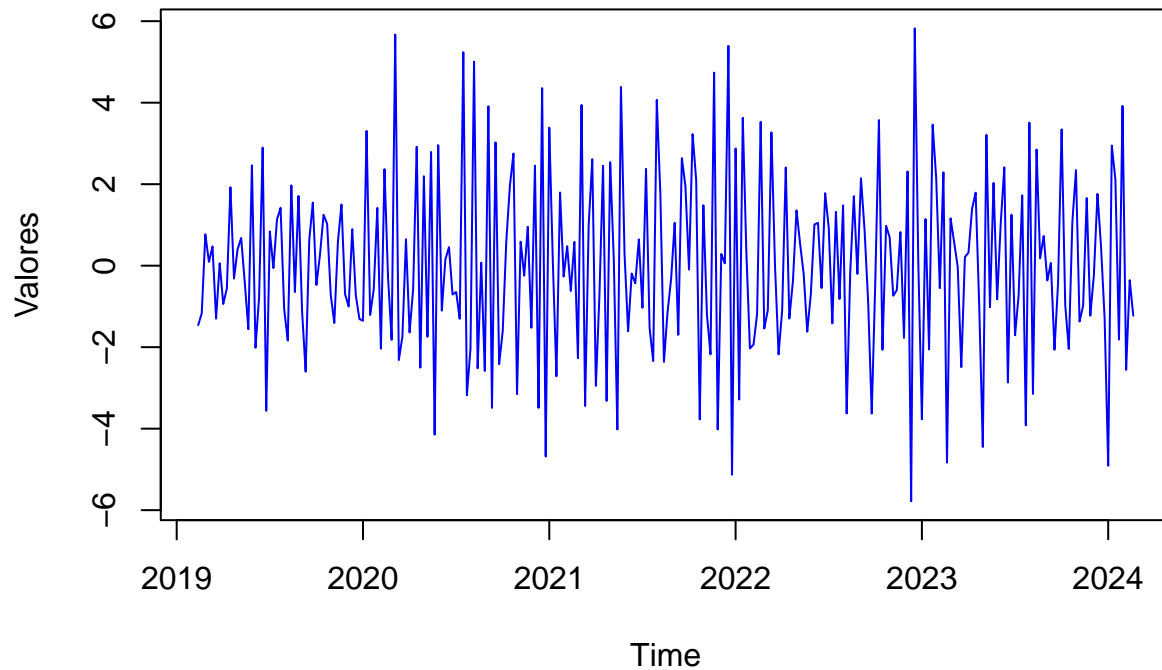
##   Indice      Fecha      Totales
## 1      1 2019-07-03 -1.45928656
## 2      2 2019-07-04 -1.16603428
## 3      3 2019-07-05  0.76901746
## 4      4 2019-07-06  0.09313228
## 5      5 2019-07-08  0.47168104
## 6      6 2019-07-09 -1.30148853
```

Serie de tiempo

```
serie_KNN_TOT_sem <- ts(data = totales_transformados$Totales,
                        start = c(2019,07,03), end = c(2024,08,05),
                        frequency = 52)
```

Gráficas de serie de tiempo

Serie semanal de productos totales



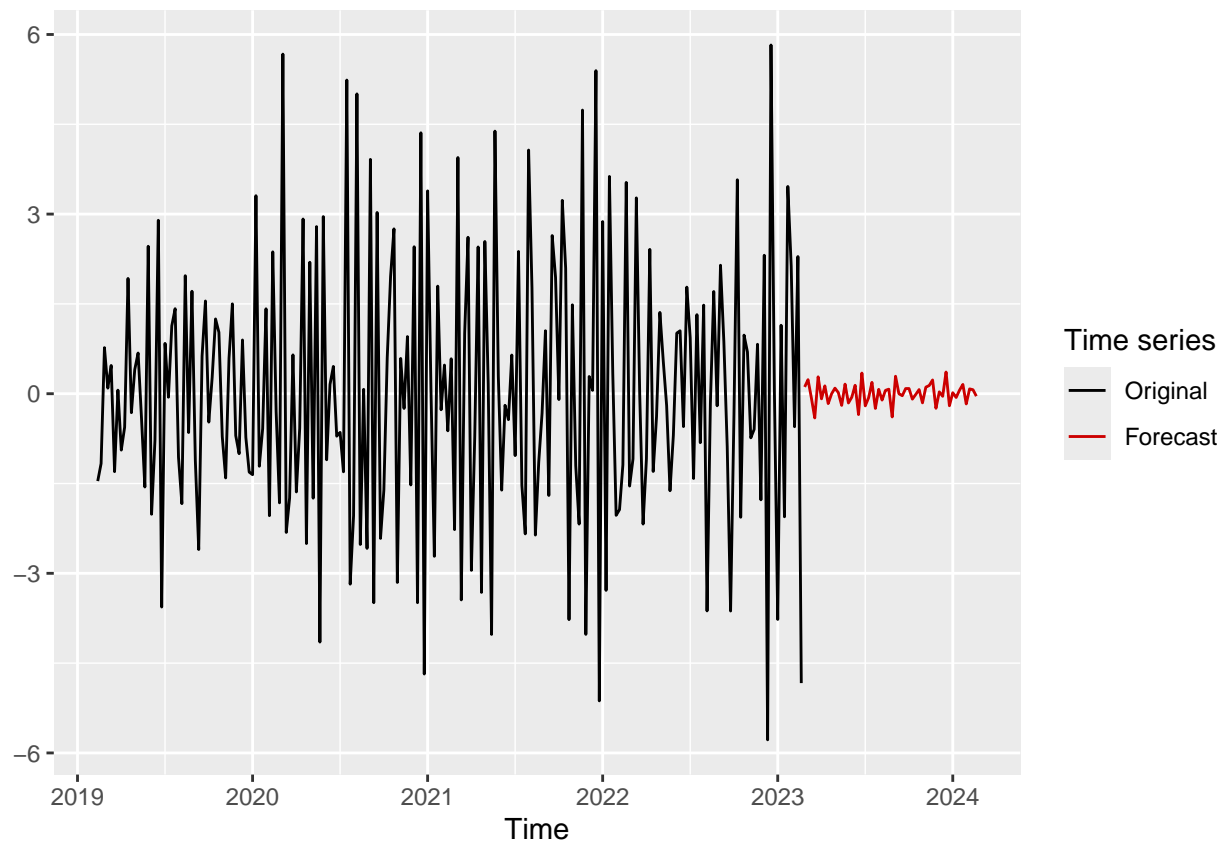
Serie de entrenamiento y prueba

```
dividida_knn_TOT_sem <- ts_split(serie_KNN_TOT_sem,  
                                sample.out = round(length(serie_KNN_TOT_sem)*0.2))  
  
train_knn_TOT_sem <- dividida_knn_TOT_sem$train  
  
test_knn_TOT_sem <- dividida_knn_TOT_sem$test
```

Modelo KNN con la estrategia de multiples entradas y multiples salidas

```
pronosticos_knn_TOT_sem <- knn_forecasting(train_knn_TOT_sem,  
                                           h = length(test_knn_TOT_sem),  
                                           lags = c(1:52),  
                                           k = 52, msas = "MIMO")
```

Gráfica pronóstico



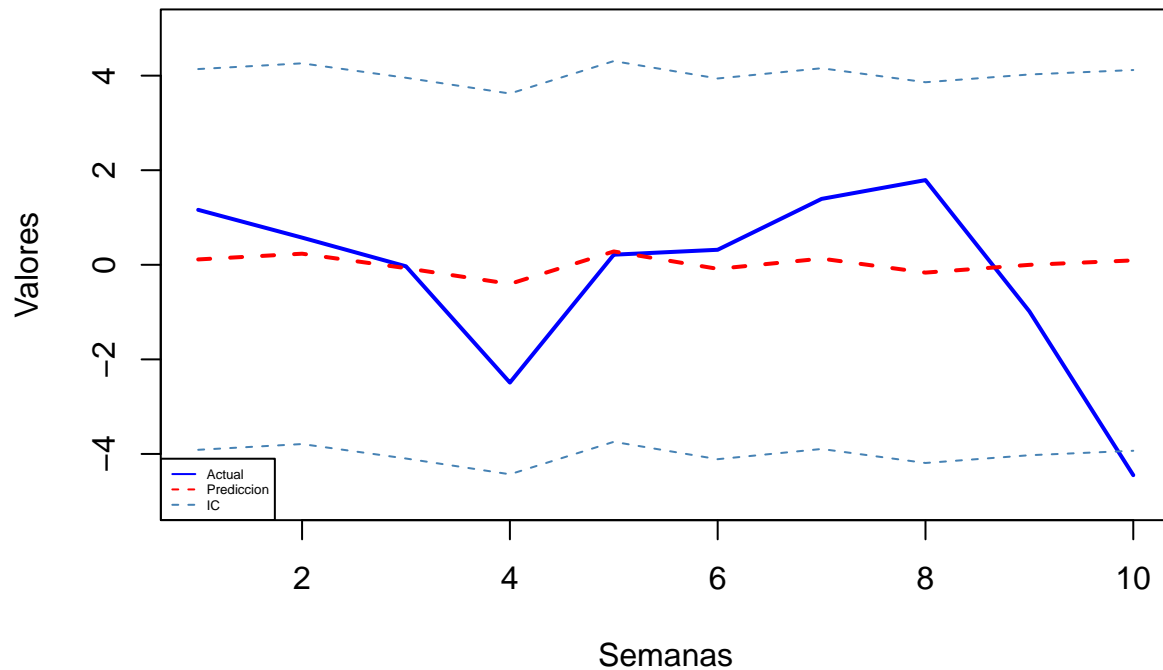
Intervalo de confianza

```
erro_TOT_knn_sem <- ( test_knn_TOT_sem - pronosticos_knn_TOT_sem$prediction)
sd_errores_TOT_knn_sem <- sd(erro_TOT_knn_sem, na.rm = T)
sd_e <- (sd_errores_TOT_knn_sem)
margen_error_TOT_knn_sem <- sd_e * 1.959964 # qnorm(0.975)

# Limites
pronosticos_knn_TOT_sem$inferior <- pronosticos_knn_TOT_sem$prediction - margen_error_TOT_knn_sem
pronosticos_knn_TOT_sem$superior <- pronosticos_knn_TOT_sem$prediction + margen_error_TOT_knn_sem
```

Gráfica con los valores de prueba y pronóstico.

Serie Prueba – Pronostico KKN semanal total



Medidas de exactitud

```
accuracy(pronosticos_knn_TOT_sem$prediction,x = test_knn_TOT_sem)
```

```
##               ME      RMSE      MAE      MPE      MAPE      ACF1 Theil's U
## Test set -0.001072913 2.034083 1.644042 90.72219 96.59325 -0.4201553 0.8671777
```

```
#               ME      RMSE      MAE      MPE      MAPE      ACF1 Theil's U
# Test set -0.001072913 2.034083 1.644042 90.72219 96.59325 -0.4201553 0.8671777
```

```
accuracy(pronosticos_knn_TOT_sem$prediction[1:10],x = test_knn_TOT_sem[1:10])
```

```
##               ME      RMSE      MAE      MPE      MAPE
## Test set -0.2631414 1.810008 1.273122 60.92667 91.45617
```

```
#               ME      RMSE      MAE      MPE      MAPE
# Test set -0.2631414 1.810008 1.273122 60.92667 91.45617
```

Se realiza el método rolling del paquete “tsfknn” que evalúa la precisión de origen rotativo de la predicción de un modelo

```
prono_knn_TOT_roll_sem <- knn_forecasting(serie_KNN_TOT_sem,
                                         h = length(test_knn_TOT_sem),
                                         lags = c(1:52),
                                         k = c(2:52), msas = "MIMO")
roll_TOT_sem <- rolling_origin(knnf = prono_knn_TOT_roll_sem,
                              h = length(test_knn_TOT_sem), rolling = T)
```

Evitando el cero absoluto

```
prueba_numeric_sem <- as.numeric(test_knn_TOT_sem)
prueba_numeric_sem[prueba_numeric_sem == 0] <- 1e-6
```

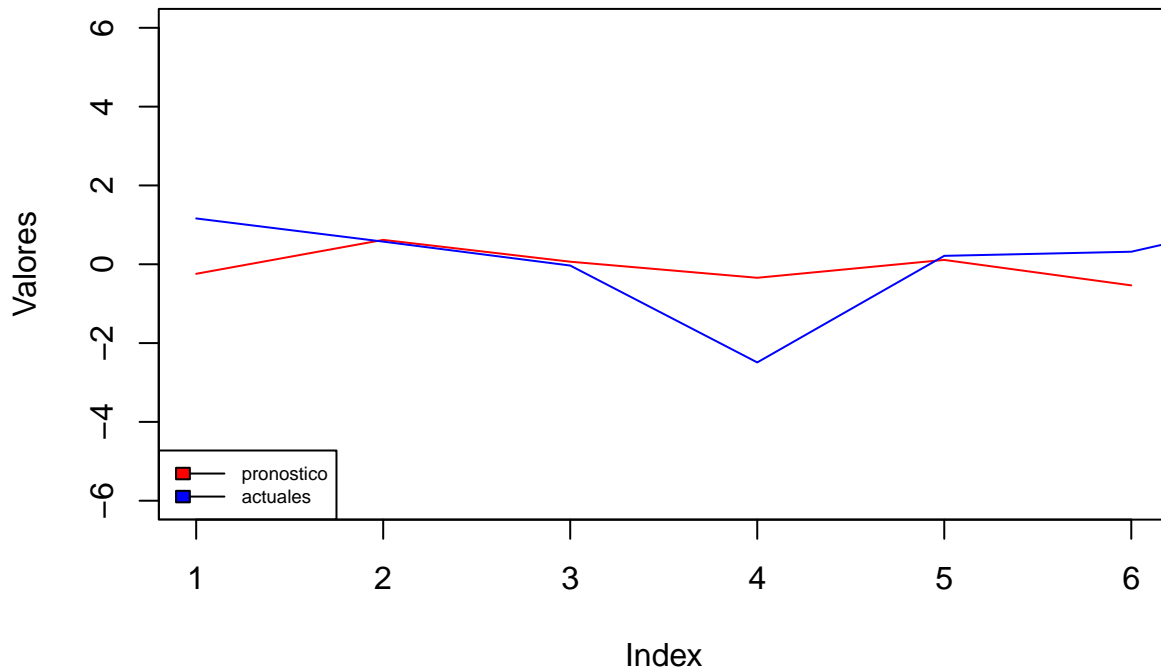
Métricas con una previsión de 10 semanas

```
accuracy(prono_knn_TOT_roll_sem$prediction[1:10],x = prueba_numeric_sem[1:10])
```

```
##           ME      RMSE      MAE      MPE      MAPE
## Test set -0.2286346 1.84927 1.335632 121.722 123.2564
```

```
#           ME      RMSE      MAE      MPE      MAPE
#  Test set -0.2286346 1.84927 1.335632 121.722 123.2564
```

Serie de pronósticos semanales de productos



Conclusiones

El modelo resultante por el método si el rolling ofrece un mejor resultado en las medidas de exactitud.