

## Uso de las series de tiempo como herramienta, para predecir comportamiento de una acción bursátil a corto plazo

M.C Armando de Jesús Ruiz Calderón. A, Mtro. Fernando Ferrero Ramón. B Dr. José de Jesús López Villalobos C.

<sup>a</sup> Instituto Tecnológico de Nuevo León, TECNM, Av Eloy Cavazos 2001, Toluca, 67170 Guadalupe, N.L. [armando.rc@nuevoleon.tecnm.mx](mailto:armando.rc@nuevoleon.tecnm.mx) Nuevo León México A .

<sup>b</sup> Instituto Tecnológico de Nuevo León, TECNM, Av Eloy Cavazos 2001, Toluca, 67170 Guadalupe, N.L. [fernando.fr@nuevoleon.tecnm.mx](mailto:fernando.fr@nuevoleon.tecnm.mx) Nuevo León México B.

<sup>c</sup> Instituto Tecnológico de Nuevo León, TECNM, Av Eloy Cavazos 2001, Toluca, 67170 Guadalupe, N.L. [jose.lv@nuevoleon.tecnm.mx](mailto:jose.lv@nuevoleon.tecnm.mx) Nuevo León México C.

### Resumen

El presente trabajo muestra la eficiencia de las series de tiempo en la predicción del comportamiento de acciones bursátiles, dado que el desconocimiento de las personas en las opciones de inversión bancaria tradicional, produce incertidumbre. El uso de las series de tiempo ayuda a dar una visión más adecuada para la toma de decisiones para invertir. Se enfoca en el uso de modelos de series temporales, utilizando el modelo ARIMA, para predecir el comportamiento a corto plazo de acciones bursátiles. Se explica detalladamente el proceso de aplicación de este modelo, desde la exploración de datos hasta la generación de predicciones, utilizando el lenguaje de programación R. Se seleccionan dos ejemplos de acciones mexicanas, Alsea y Bimbo, para demostrar la utilidad del modelo ARIMA. Los resultados muestran predicciones acertadas en ambas acciones, confirmando la eficacia de las series temporales en la toma de decisiones de inversión. Se destaca la importancia de esta herramienta para personas sin conocimientos financieros profundos. Además, se sugiere la implementación de un modelo semi supervisado de aprendizaje, como trabajo futuro.

**Palabras clave**—Acciones bursátiles, predicción de comportamiento, Series de tiempo.

### Abstract

*The present work shows the efficiency of time series in predicting the behavior of stock market actions, given that people's ignorance of traditional banking investment options produces uncertainty. The use of time series helps to provide a vision more appropriate for making investment decisions. It focuses on the use of time series models, using the ARIMA model, to predict the short-term behavior of stock stocks. The process of applying this model is explained in detail, from data exploration to the generation of predictions, using the R programming language. Two examples of Mexican stocks, Alsea and Bimbo, are selected to demonstrate the usefulness of the ARIMA model. The results show accurate predictions in both stocks, confirming the effectiveness of time series in making investment decisions. The importance of this tool for people without deep financial knowledge is highlighted.*

*Furthermore, the implementation of a semi-supervised learning model is suggested as future work..*

**Keywords**— Maximum of 5 keywords in English alphabetical order.

### 1. INTRODUCCIÓN

En años recientes, el aumento en la preocupación de las personas por generar o incrementar un ahorro para el retiro, y las malas condiciones bancarias tradicionales, provocan que las personas con posibilidades de ahorro, y las generaciones de jóvenes busquen nuevas estrategias para incrementar su capital; entre estas estrategias, están las diferentes formas de invertir, al existir un portafolio amplio de inversiones, muchas personas, caen confundidos por el tipo de inversión, y otras más, invierten sin tener los conocimientos necesarios y en muchos casos, el resultado no es el esperado.

De acuerdo con Kar Wai Lim [1] Los modelos de series de tiempo han sido la base para cualquier estudio de un comportamiento o proceso durante un período de tiempo; esto aunado a Petrov [2] quien señala que el uso de las series de tiempo, se pueden aplicar en diversas áreas de la investigación, e industrias tal como las aplicaciones que producen y tratan dichos datos son datos médicos, servicios basados en la ubicación, servicios inalámbricos redes de sensores, [8], además de ser un área de estudio de gran dinamismo para las finanzas [4,5,6,7,9], por nombrar solo algunos, vuelve a las series de tiempo el algoritmo idóneo para este trabajo.

Las series temporales o series de tiempo, son conjuntos de datos recopilados o registrados en intervalos de tiempo regulares. Estos datos pueden representar observaciones de una variable en función del tiempo y se utilizan para analizar patrones temporales, identificar tendencias y hacer predicciones sobre el comportamiento futuro. Las series de tiempo son comunes en diversos campos, como ya se ha mencionado en finanzas, economía, meteorología, epidemiología, entre otros.

Características de las series de tiempo:

1. Orden temporal: Los datos están organizados en orden cronológico, y la secuencia temporal es esencial para comprender el comportamiento de la variable a lo largo del tiempo.
2. Dependencia temporal: Las observaciones en un momento dado pueden depender de las observaciones anteriores. Esto implica que el valor actual de la serie puede estar relacionado con los valores pasados.

Las predicciones en las series tiempo son importantes para anticipar el comportamiento futuro de la variable de interés. Algunas de las técnicas comunes utilizadas para hacer predicciones en series temporales incluyen:

- Modelos autorregresivos (AR): Estos modelos asumen que el valor actual de la serie está relacionado con sus valores pasados. Un modelo AR(p) utiliza los p valores más recientes para predecir el siguiente.
- Modelos de medias móviles (MA): Estos modelos asumen que el valor actual de la serie está relacionado con los errores residuales pasados. Un modelo MA(q) utiliza los q errores residuales más recientes para hacer predicciones.
- Modelos autorregresivos de medias móviles (ARIMA): La combinación de los modelos AR y MA en un solo marco se conoce como ARIMA. Un modelo ARIMA(p, d, q) tiene componentes autorregresivos, diferenciación y de medias móviles.
- Modelos estacionales: Incorporan patrones estacionales en las predicciones para manejar variaciones recurrentes que ocurren en ciertos momentos del año.
- Modelos de suavizado exponencial: Estos modelos son útiles para series temporales con tendencias y/o estacionalidad y se basan en promedios ponderados de observaciones pasadas.
  - Modelos de regresión: Se utilizan cuando la serie temporal está influenciada por variables externas.

Para aplicar predicciones en una serie de tiempo, generalmente se sigue este proceso:

- a) Exploración de datos: Comprender la estructura de la serie, identificar tendencias, estacionalidades y posibles patrones.
- b) Selección del modelo: Elegir un modelo adecuado basado en la exploración de datos y diagnósticos del comportamiento temporal.
- c) Entrenamiento del modelo: Ajustar el modelo a los datos históricos para aprender los parámetros y comportamientos del pasado.
- d) Validación del modelo: Evaluar el rendimiento del modelo utilizando datos de prueba para asegurarse de que pueda generalizar bien a nuevas observaciones.
- e) Generación de predicciones: Utilizar el modelo entrenado para hacer predicciones futuras.
- f) Evaluación de predicciones: Comparar las predicciones con los datos reales para evaluar la precisión del modelo.

Este proceso puede repetirse y ajustarse según sea necesario para mejorar la precisión de las predicciones a medida que se obtienen más datos.

El modelo más adecuado para este trabajo es ARIMA (Autorregresivo Integrado de Medias Móviles) el cual es una técnica de modelado estadístico utilizada para analizar y predecir series temporales. Este modelo combina componentes autorregresivos (AR), de integración (I) y de medias móviles (MA).

- Autorregresivo (AR): Refleja la relación entre una observación actual y sus observaciones pasadas. En un modelo AR(p), se utilizan los p valores más recientes para predecir el siguiente.

- Integrado (I): Representa la diferenciación de la serie temporal para hacerla estacionaria. La diferenciación implica restar una observación de otra para eliminar tendencias o patrones no estacionarios. La cantidad de diferenciación necesaria se denota como "d" en un modelo ARIMA(p, d, q).
- Medias Móviles (MA): Describe la relación entre una observación actual y un término de error residual de observaciones pasadas. En un modelo MA(q), se utilizan los q errores residuales más recientes para hacer predicciones.

La notación general de un modelo ARIMA se expresa como ARIMA(p, d, q). El orden de los componentes p, d y q se determina mediante análisis exploratorio de datos y diagnósticos del comportamiento temporal de la serie. El proceso para calcular un modelo ARIMA incluye los siguientes pasos:

1. Exploración de datos: Comprender la serie temporal, identificar tendencias, estacionalidades y patrones.
2. Estacionalización: Si la serie no es estacionaria, aplicar diferenciación para hacerla estacionaria. El valor de d se determina por el número de diferenciaciones necesarias.
3. Identificación de componentes ARIMA: Utilizar gráficos de autocorrelación y autocorrelación parcial para identificar posibles valores de p y q.
4. Selección de modelo: Utilizar métodos como el criterio de información de Akaike (AIC) o el criterio de información bayesiano (BIC) para seleccionar el mejor modelo ARIMA.
5. Entrenamiento del modelo: Ajustar el modelo ARIMA a los datos de entrenamiento utilizando técnicas de estimación de parámetros.
6. Validación del modelo: Evaluar el rendimiento del modelo utilizando datos de prueba para asegurarse de que pueda generalizar bien a nuevas observaciones.
7. Generación de predicciones: Aplicar el modelo entrenado para hacer predicciones futuras.

Para el presente trabajo, se desarrolló en lenguaje R dada su naturaleza y por tener las herramientas necesarias para el análisis de datos.

En R, para poder utilizar el modelo de ARIMA, se debe tener instalada e inicializada la biblioteca "forecast" la cual es comúnmente utilizada para trabajar con modelos de análisis de datos. La función "auto.arima()" dentro de este paquete es requerida para seleccionar automáticamente los órdenes p, d y q.

```
Código en R  
library(forecast)
```

```
# Crear una serie temporal (reemplazar con tus propios datos)
serie_temporal <- ts(c(x1,...,xn))
# Seleccionar automáticamente los órdenes p, d, y q
modelo_arima <-
auto.arima(serie_temporal)
# Imprimir el modelo seleccionado
print(modelo_arima)
# Hacer predicciones
predicciones <- forecast(modelo_arima, h
= 5) #h representa la frecuencia en horas
print(predicciones)
```

Este código ilustra de forma genérica cómo utilizar “*auto.arima()*” para seleccionar automáticamente los órdenes del modelo ARIMA y cómo hacer predicciones con el modelo ajustado.

En concreto el algoritmo ARIMA es especialmente útil para conjuntos de datos que pueden ser mapeados a series de tiempo fijas. Las propiedades estadísticas de las series temporales fijas, como las autocorrelaciones, son independientes del tiempo

Por otro lado, una acción bursátil es un activo financiero que representa una parte del capital social de una sociedad anónima. La adquisición de acciones otorga al inversionista ciertos derechos sobre la empresa y obtiene la categoría de socio. En consecuencia, el comprador de una acción o acciones se convierte en dueño de la empresa considerando la proporción de acciones compradas.[3]

## 2 OBJETIVO

Mostrar la utilidad del uso de las series de tiempo en la predicción del comportamiento de una acción bursátil a corto plazo.

## 3 JUSTIFICACIÓN

La falta de conocimientos financieros, así como el desconocimiento en el uso de las herramientas financieras existentes genera un vacío en el entendimiento de las formas de inversión, lo que provoca errores al momento de tomar decisiones; el uso de series de tiempo con lenguaje R facilita el entendimiento del comportamiento de una acción bursátil a corto plazo, lo que ayuda a tener certeza en la toma de decisiones

## 4 CONTRIBUCIÓN

Mostrar la utilidad del uso de las series de tiempo para conocer el posible comportamiento a corto plazo de una acción bursátil, y que sea de utilidad para cualquier persona con un poco de conocimiento en programación, y pueda tomar mejores decisiones para invertir.

## 5 METODOLOGIA

Para el presente trabajo se seleccionaron dos acciones bursátiles de dos segmentos de mercado diferentes, y en el que ambas empresas son de una volatilidad baja. Ambas empresas son mexicanas la primera pertenece al sector de servicios y

bienes de consumo no básico; al subsector de servicios al consumidor y al ramo de hoteles, restaurantes y esparcimiento; así mismo la segunda acción estudiada, pertenece al segmento de productos de consumo frecuente y al subsector alimentos, bebidas y tabaco en el ramo de alimentos.

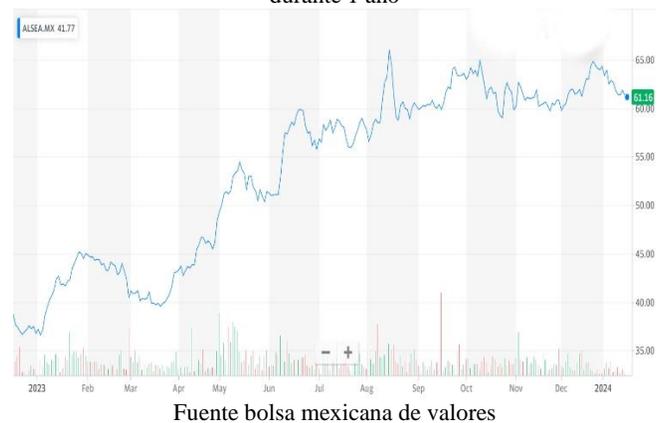
- 1.- Se realizó la selección de las acciones a estudiar.
- 2.-De los repositorios de datos de la bolsa mexicana de valores se obtuvieron los costos de las acciones seleccionadas del último año [6,7].
- 3.-Se realiza el desarrollo del algoritmo ARIMA en lenguaje R..
- 4.- Con base en el formato que pide ARIMA, se ordenan los datos obtenidos de las acciones a estudiar para su procesamiento.
- 5.- Se generan los archivos .csv, para ser procesados por las bibliotecas del lenguaje R.
- 6.-Se realizan las pruebas de las dos acciones.

## 6 RESULTADOS

Se realizó el análisis de los datos de ambas acciones, del período comprendido de enero del 2023 a enero del 2024; obteniéndose los siguientes resultados:

- 1.- Se observa que la acción de Alsea tiende a aumentar de precio, la gráfica de la figura (1), muestra los valores diarios durante un año de enero del 2023 a enero del 2024; aun cuando esa acción presenta un comportamiento un poco volátil, debido a sus variaciones diarias en su valor, se puede observar la tendencia de aumento en su valor.

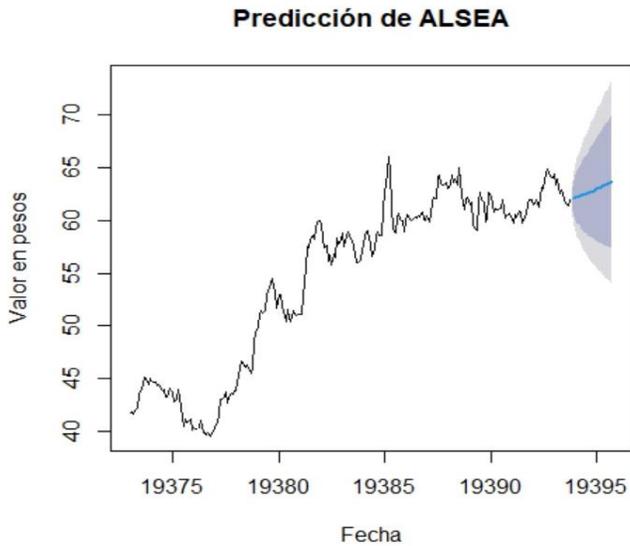
Fig. 1. Gráfica que muestra la cotización de la acción de Alsea durante 1 año



Fuente bolsa mexicana de valores

En la figura, (2), se muestra la gráfica que se obtiene después de ejecutar el algoritmo ARIMA, para el análisis de datos, obteniéndose como resultado, que el valor de la acción el tiende a aumentar.

Fig. 2. La gráfica muestra la predicción para Alsea



Fuente: elaboración propia a partir de la ejecución del algoritmo

La figura.2 muestra la predicción del comportamiento de la acción de Alsea realizada por ARIMA, donde se muestra la tendencia a aumentar su valor y considerando que el valor de la acción fue de \$61.78 MXN y el valor al momento de revisar la cotización era de \$61.89 MXN, que se muestra en la fig. 3., lo que refleja un alza en el valor de cada acción lo que confirma la utilidad de las series de tiempo

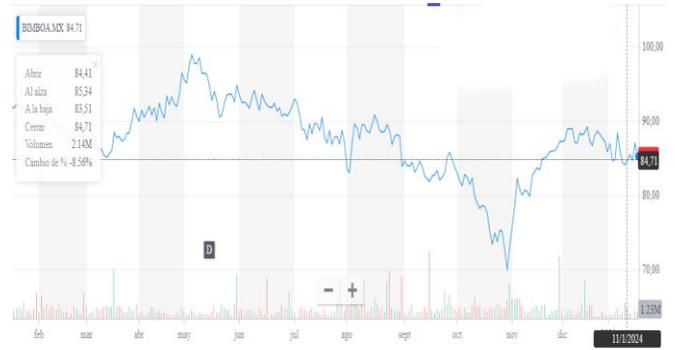
Fig. 3. Cotización de Alsea el 17 de enero 2024



Fuente bolsa mexicana de valores

2.- La figura (4) muestra los valores de la acción de Bimbo desde enero 2023 hasta el 11 enero del 2024, donde se puede observar una tendencia a la baja, posteriormente una caída en el precio y una ligera recuperación, pero con una tendencia a la baja lo que se ve reflejado en la predicción.

Figura 5 cotización de Bimbo de enero del 2023 al 11 de enero del 2024.

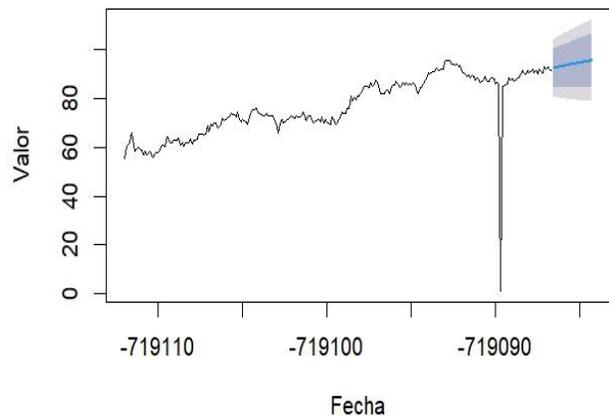


Fuente Bolsa mexicana de valores

La figura 5 muestra la predicción del comportamiento de la acción de bimbo en donde se observa que a lo largo del año fue bajando su valor tuvo un desplome en el valor y nuevamente repuntó, con una ligera tendencia a mantener en su valor.

Fig. 5. Muestra el comportamiento y la predicción del valor de la acción

### Predicción Bimbo



Fuente: elaboración propia a partir de la ejecución del algoritmo

La figura (6) muestra la predicción que el algoritmo ARIMA realiza del comportamiento de la acción de Bimbo, donde se muestra la tendencia a aumentar su valor y señalando que el valor de la acción el 11 de enero fue de \$84.71 MXN y el valor al momento de revisar la cotización era de \$85.07 MXN, que se muestra en la figura (6), lo que refleja un alza en el valor de cada acción lo que vuelve a confirmar la utilidad de las series de tiempo.

Estos resultados nos muestran la eficiencia del uso de las series de tiempo, el algoritmo de ARIMA, tiene parámetros de ajuste, de donde se utilizó como tiempo en días 5 para realizar las predicciones

Fig. 6. Costo de la acción de Bimbo el 17 de enero 2024



Es importante resaltar que no es eficiente realizar un cálculo de la pendiente o regresión lineal dentro del período a estudiar, puesto que no refleja realmente el comportamiento de cada acción, además de que la no linealidad del comportamiento bursátil, debido a la volatilidad del mercado, y considerando que la regresión lineal asume que los residuos, son independientes e idénticamente distribuidos. Estas suposiciones pueden no ser válidas en el contexto de los datos bursátiles, donde la relación entre los precios de las acciones y otros factores puede no ser lineal y los residuos pueden mostrar patrones de autocorrelación también la regresión lineal, no considera la estocasticidad de los datos, pues los precios de las acciones están sujetos a fluctuaciones aleatorias en el tiempo, y por último con la regresión lineal no se tiene sensibilidad a puntos atípicos, y a que puede ser altamente sensible a valores extremos en los datos lo que puede distorsionar los resultados y conducir a conclusiones erróneas.

## 6. CONCLUSIONES

Considerando las 2 acciones estudiadas y los resultados mostrados podemos concluir que:

- 1.- Las series de tiempo son una herramienta útil para conocer el posible comportamiento de una acción, bursátil a corto plazo, y con la configuración utilizada es para no más de 5 días.
- 2.-El uso de las series de tiempo, para personas que no son del área de finanzas resulta de utilidad para una toma de decisiones.
- 3.-El uso del modelo ARIMA (Autoregressive Integrated Moving Average) es una herramienta fundamental para el análisis de series de tiempo dada su flexibilidad pues ARIMA es altamente flexible y puede adaptarse a una amplia gama de patrones y estructuras en los datos de series de tiempo. Puede manejar tendencias, estacionalidad y otros tipos de comportamientos complejos presentes en los datos; además tiene gran capacidad para capturar la dependencia temporal. ARIMA es capaz de capturar la dependencia temporal en los datos, lo que significa que puede modelar cómo los valores pasados de una serie de tiempo influyen en los valores futuros. Esto lo logra mediante los componentes autorregresivos y de medias móviles del modelo.

4. Este algoritmo es muy adecuado para el modelado de estacionalidad, pues puede incorporar componentes estacionales en el modelo, lo que es crucial para el análisis de datos que exhiben patrones estacionales recurrentes, como las ventas minoristas, la demanda de energía o los datos financieros.

5. ARIMA tiene una gran capacidad para manejar estacionariedad y puede manejar series de tiempo que no son estacionarias mediante la diferenciación, lo que implica tomar diferencias entre observaciones sucesivas para hacer que la serie de tiempo sea estacionaria. Esto es útil porque muchos procesos del mundo real no son estacionarios y requieren este tipo de transformación para el análisis.

6. ARIMA es ampliamente utilizado para la predicción de valores futuros en series de tiempo. Una vez que se ha ajustado a los datos históricos, se puede utilizar para hacer pronósticos sobre los valores futuros de la serie de tiempo con un alto grado de confianza.

7.- Se debe ajustar el algoritmo utilizado a un modelo semi supervisado de aprendizaje para mejorar la eficiencia tal como lo plantea Gonzalez [10]

Como trabajo futuro se debe realizar una aplicación, ya sea para móvil o para escritorio en la que se pueda incorporar los datos a analizar y se realice de forma automática, y que sea de uso sencillo e intuitivo

## 7. REFERENCIAS

- [1] Kar War Lim et.al, "Bayesian Bivariate Hawkes", TSAA '16: Workshop on Time Series Analytics and Applications Hobart TAS Australia 6 December 2016, ISBN 978-1-4503-4820-1
- [2] Daniel Petrov, Rakan Alseghayer, Mohamed Sharaf, Panos K. Chrysanthis, Alexandros Labrinidis. 2017. Interactive Exploration of Correlated Time Series. In Proceedings of ExploreDB'17, Chicago, IL, USA, 2017, 6 pages. DOI: <http://dx.doi.org/10.1145/3077331.3077335>
- [3] <https://marketdata.com.py/educacion/conceptos/que-es-una-accion-bursatil-43063/>
- [4] <https://www.bmv.com.mx/es/emisoras/perfil/BIMBO-5163>
- [5] <https://www.bmv.com.mx/es/emisoras/perfil/ALSEA-5059>
- [6] <https://finance.yahoo.com/quote/BIMBOA.MX/>
- [7] <https://finance.yahoo.com/chart/ALSEA.MX/>
- [8] Reynold Cheng, Dmitri V. Kalashnikov, and Sunil Prabhakar. 2003. Evaluating Probabilistic series over Imprecise Data ACM SIGMOD'03, 551–562
- [9] 2] Krisztian Buza, Alexandros Nanopoulos, and Lars Schmidt-Thieme. 2010. Timeseries classification based on individualised error prediction. In 13th International Conference on Computational Science and Engineering. IEEE, 48–54

[10] González Castellanos, Mabel & Bergmeir, Christoph & Triguero, Isaac & Rodríguez, Yanet & Benítez, José. (2018). Self-labeling techniques for semi-supervised time series classification: an empirical study. *Knowledge and Information Systems*. 55. 10.1007/s10115-017-1090-9.