

# APLICACIÓN DEL PROCESO KDD EN LA COMPRA DE ACCIONES DE GRUPO BIMBO

## MINERÍA DE DATOS

**González Flores Diana Patricia**

**Matrícula:** 170812

8° Semestre

Ingeniería en Tecnologías de la Información

**Docente:** Manuel Chávez Pérez



## Contenido

Resumen .....	2
Marco teórico.....	2
Desarrollo .....	4
Integración y recopilación .....	4
Series de datos utilizadas .....	4
Variables consideradas .....	4
Tipos de dato .....	5
Justificación de las series de datos utilizadas .....	5
Creación del conjunto de datos.....	6
Selección, limpieza y transformación.....	10
Estrategia para datos incompletos .....	11
Estrategia para considerar los atributos relevantes .....	11
Minería de datos .....	13
Tarea de minería de datos .....	13
Tipo de modelo .....	13
Algoritmo .....	13
Ejecución de código en R.....	13
Evaluación e interpretación.....	17
Técnicas de evaluación.....	17
Medidas de evaluación.....	17
Modelo final construido .....	17
Interpretación y contextualización .....	19
Difusión, uso y monitorización .....	20
6 de Mayo de 2021 .....	20
7 de Mayo de 2021 .....	26
Resultados .....	30
Diferencias entre el precio de cierre pronosticado y el valor real.....	30
Ganancias y Pérdidas.....	30
Conclusiones.....	32
Propuestas .....	33
Código en R .....	35
Bibliografía .....	37

## Resumen

Día con día, y más con la evolución tecnológica, una gran cantidad de datos se encuentra a nuestro alrededor, desde registros de personas, datos de alguna enfermedad, registros de internet, datos de movilidad, compras, valores de empresas y una infinidad más.

Estos datos crecen sin parar, por lo que es interesante saber qué más podemos aprender de esta gran cantidad de datos. Aquí entra el concepto de “Minería de datos”, que se refiere al análisis de grandes cantidades de datos para encontrar patrones y llegar al conocimiento.

Entonces, si se tienen grandes cantidades de datos, éstos se deben aprovechar para nuestro beneficio y el de los demás, ya que por ejemplo, se podría ayudar en la investigación de una nueva enfermedad, para el análisis de la seguridad de una ciudad, para aumentar las ventas de una empresa, reducir riesgos de una organización, incrementar ingresos, entre otras implicaciones más.

Debido a lo anterior, en el presente proyecto de la materia Minería de Datos de la Universidad Politécnica de San Luis Potosí se utilizarán técnicas de minería de datos, como redes neuronales y el algoritmo Backpropagation a través del proceso KDD, en la compra de acciones en la Bolsa Mexicana de Valores en la plataforma AcciTrade Coach para decidir en qué momento es el más apropiado para comprar acciones y así obtener ganancias del dinero invertido.

## Marco teórico

Para el desarrollo de este proyecto, se deben tener en claro ciertos conceptos, los cuales son:

### **Dato**

Representación de una variable, indica el valor asignado a una cosa. Un dato no contiene información relevante, solo se puede apreciar información relevante al analizar un conjunto de datos.

### **Minería de datos**

Como menciona Molina Félix (2001), la minería de datos se refiere a “la integración de un conjunto de áreas que tienen como propósito la identificación de un conocimiento obtenido a partir de las bases de datos que aporten un sesgo hacia la toma de decisión”. Es una etapa del proceso KDD (Knowledge Discovery in Databases), en español, descubrimiento de conocimiento en bases de datos, donde se buscan patrones o modelos a partir de los datos recopilados.

### **Proceso KDD**

Knowledge Discovery in Databases, en español, descubrimiento de conocimiento en bases de datos. Molina Félix(2001) define este proceso como “ la extracción no trivial de información potencialmente útil a partir de un gran volumen de datos, en el cual la información está implícita, donde se trata de interpretar grandes cantidades de datos y encontrar relaciones o patrones, para conseguirlo harán falta técnicas de aprendizaje, estadística y bases de datos”. Por lo que se refiere al proceso iterativo e interactivo de descubrir patrones en datos, con el objetivo de llegar al conocimiento. Este proceso cuenta con cinco fases las cuales son:

1. Integración y recopilación
2. Selección, limpieza y transformación
3. Minería de datos
4. Evaluación e interpretación
5. Difusión y uso

### **Redes neuronales**

Técnica basada en modelos computacionales del aprendizaje humano a través de las neuronas del cerebro, como menciona Beltrán Martínez (2018) , “Las redes neuronales son una nueva forma de analizar la información con una diferencia fundamental con respecto a las técnicas tradicionales: son capaces de detectar y aprender patrones y características dentro de los datos.”, estas redes se comportan como nuestro cerebro, pues aprenden de la experiencia y del pasado, para posteriormente aplicar ese conocimiento para resolver nuevos problemas.

### **Algoritmo Backpropagation**

Método de aprendizaje supervisado de entrenamiento para redes neuronales, con dos fases: primero se aplica un patrón de entrada, el cual se propaga por las distintas capas que componen la red hasta producir la salida de la misma. Esta salida se compara con la salida deseada y se calcula el error cometido por cada neurona de salida. Estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de las capas intermedias. Cada neurona recibe un error que es proporcional a su contribución sobre el error total de la red. Basándose en el error recibido, se ajustan los errores de los pesos de cada neurona. (Toral Barrera, 2009)

### **Bolsa de Valores**

Roldan (2020) la define como un mercado donde se contactan demandantes y oferentes de capital, estos realizan transacciones a través de intermediarios autorizados. Aquí las empresas que requieren financiamiento y ahorradores (personas u organizaciones) negocian e intercambian capital, ya que los ahorradores buscan invertir su dinero con el fin de obtener rentabilidad. Normalmente se negocia renta variable (acciones) o renta fija (deuda), la renta variable es más popular y conocida.

## Acción

Activo financiero, representa una parte proporcional del capital social de una sociedad anónima (S.A.), cuando se adquieren acciones, el comprador recibe derechos sobre la empresa y se convierte en socio de la misma, por lo que el comprador se convierte en dueño de la empresa en proporción a las acciones compradas. (Jiménez Bermejo, 2015)

## AcciTrade Coach

Plataforma de educación financiera y entretenimiento de AcciTrade, con la cual se puede aprender a invertir en la Bolsa Mexicana de Valores. En esta plataforma se puede comprar y vender acciones, operar Fondos de Inversión Citibanamex, utilizando dinero virtual dentro del simulador.

## Desarrollo

### Integración y recopilación

**Objetivo:** Determinación de las fuentes de información, para posteriormente transformar los datos a un formato común para unificar toda la información.

### Series de datos utilizadas

Para predecir el valor del precio de las acciones de grupo bimbo (BIMBOA.MX) se plantea usar los datos disponibles en Yahoo Finanzas.

Además de los datos de Grupo Bimbo, se usarán los datos de otras variables. Las series de datos que se utilizarán son:

El precio del dólar

- USD/MXN (MXN=X)

IPC México (Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores)

- IPC MEXICO (^MXX)

Materias primas

- Petróleo, Crude Oil Jun 21 (CL=F)
- Maíz, Corn Futures,Jul-2021 (ZC=F)
- Avena, Oat Futures,Jul-2021 (ZO=F)
- Trigo, KC HRW Wheat Futures,Jul-2021 (KE=F)
- Azúcar, Sugar #11 Jul 21 (SB=F)

### Variables consideradas

Open: precio de la acción al abrir la bolsa

High: el precio más alto alcanzado

Low: el precio más bajo alcanzado

Close: precio de la acción al cerrar la bolsa

Adj Close: precio al cerrar la bolsa, dividendos y división de acciones

Volume: total de acciones compradas y vendidas

De BIMBOA.MX, los datos utilizados son:

- Open
- High
- Low
- Close
- Adj Close
- Volume

Del precio del dólar, IPC y materias primas sólo se consideró el precio, por lo que se tomaron los datos de :

- High
- Low

### Tipos de dato

BIMBOA.MX : Peso mexicano MXN

MXN=X: Peso mexicano MXN

IPC MEXICO (^MXX) : Peso mexicano MXN

Crude Oil Jun 21 (CL=F): Dólar estadounidense USD

Maíz, Corn Futures,Jul-2021 (ZC=F): USX

Avena, Oat Futures,Jul-2021 (ZO=F): USX

KC HRW Wheat Futures,Jul-2021 (KE=F): USX

Sugar #11 Jul 21 (SB=F): USX

### Justificación de las series de datos utilizadas

#### **BIMBOA.MX**

Se utilizarán los datos de esta serie ya que se pronosticará el precio de estas acciones, así que es necesario tomar en cuenta los datos relacionados directamente con el precio, como el precio al abrir la bolsa de valores, los precios máximos y mínimos, etc., pues se planea predecir el precio de cierre de las acciones (close) para tener todo el día financiero para tomar decisiones acerca de comprar o vender.

#### **Precio del dólar USD/MXN (MXN=X)**

El precio del dólar puede afectar los precios, sobre todo en productos importados. También puede afectar las exportaciones, incluso otros sectores como el turismo.

Específicamente, Grupo Bimbo tiene presencia en 33 países en América, Europa, África y Asia, por lo que se define como una empresa de consumo global. El precio del dólar puede afectar los precios de las materias primas que utiliza, del combustible e incluso las exportaciones de sus productos a otros países, por lo que es importante considerar esta serie de datos ya que puede afectar el precio de sus acciones.

### **IPC México (^MXX)**

El IPC es el índice bursátil principal de México, proviene de las siglas de índice de precios y cotizaciones de México. Este índice se construye con una muestra de las empresas emisoras listadas en la bolsa de valores de México (BMV), esta muestra consiste en las 35 empresas más importantes del país. Grupo Bimbo forma parte de estas empresas.

Este índice refleja cuanto ganó o perdió la Bolsa Mexicana de Valores (BMV) en una jornada.

El valor del IPC puede afectar en diversos factores de la economía, ya que puede hacer que suban o bajen los precios, lo que puede afectar el número de ventas de una empresa, afectar el poder adquisitivo de las personas, etc. Debido a esto se considera que este índice puede afectar el precio de las acciones de BIMBOA.MX.

### **Materias Primas**

Grupo Bimbo se define como “Empresa de consumo global líder en la industria de la panificación y un jugador relevante en snacks.”

En esta empresa se producen y distribuyen pan de caja fresco y congelado, bollos, galletas, pastelitos, english muffins, bagels, productos empacados, tortillas, botanas saladas y confitería, entre otros.

Para lograr lo anterior se necesitan diversas materias primas, por lo que en este caso se consideran maíz, avena, trigo y azúcar, ya que se ven involucradas en la producción de panificación y snacks. Existe un mercado de materias primas que cotizan en la bolsa de valores, por lo que se consideraron sus precios ya que estos pueden afectar en las producciones de Grupo Bimbo y por ende afectar el precio de sus acciones.

Grupo Bimbo menciona en su sitio web que cuentan con una de las redes de distribución más grandes del mundo, por lo que se necesitan diversos medios de transporte que utilizan diversos combustibles como diésel y gasolina, por lo que se considera el precio del petróleo.

### **Creación del conjunto de datos**

1. Ingresar a <https://es-us.finanzas.yahoo.com/>

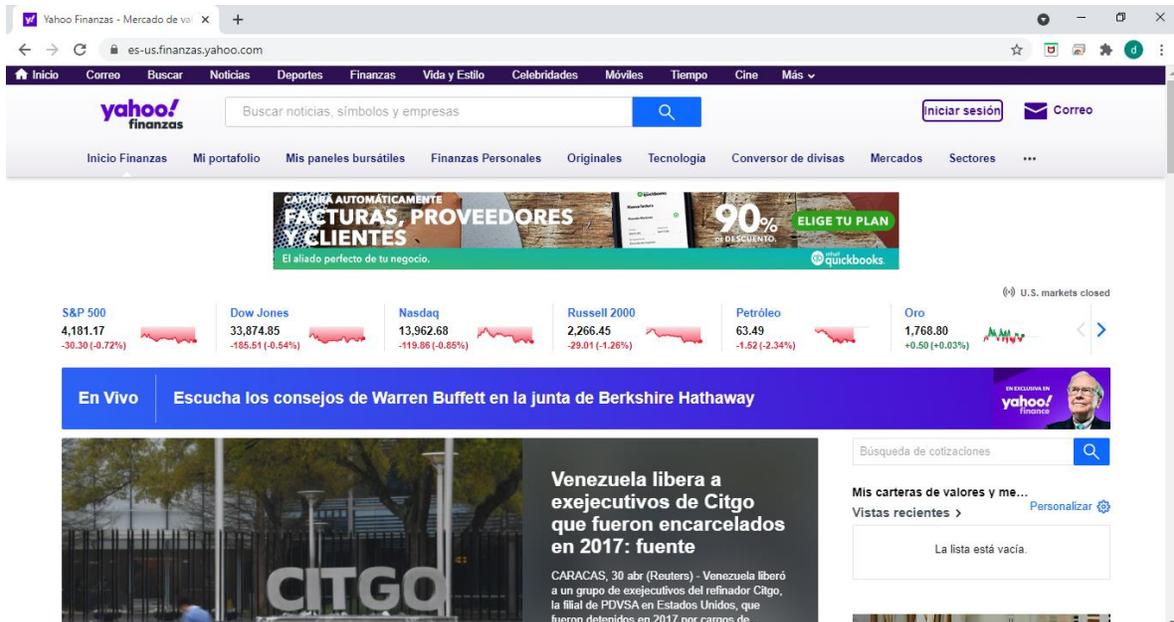


Ilustración 1 Sitio Yahoo Finanzas

2. Ingresar el nombre de la empresa en el buscador y dar clic en el resultado.

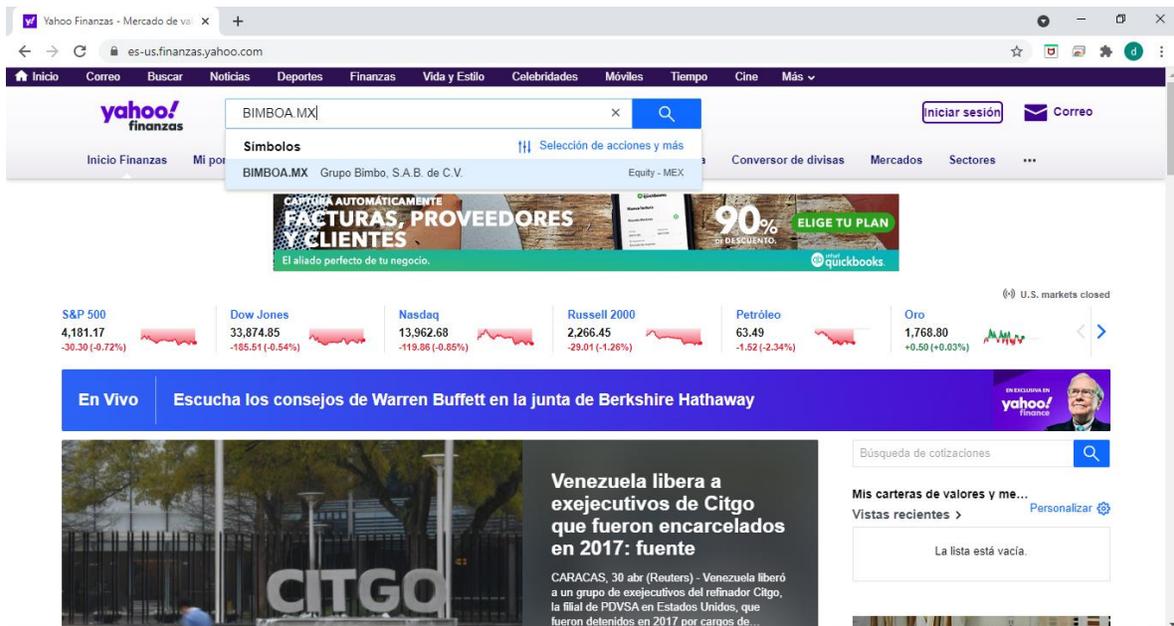


Ilustración 2 Búsqueda de la empresa

3. Se mostrará el resumen de los precios de las acciones de la empresa, dar clic en datos históricos.

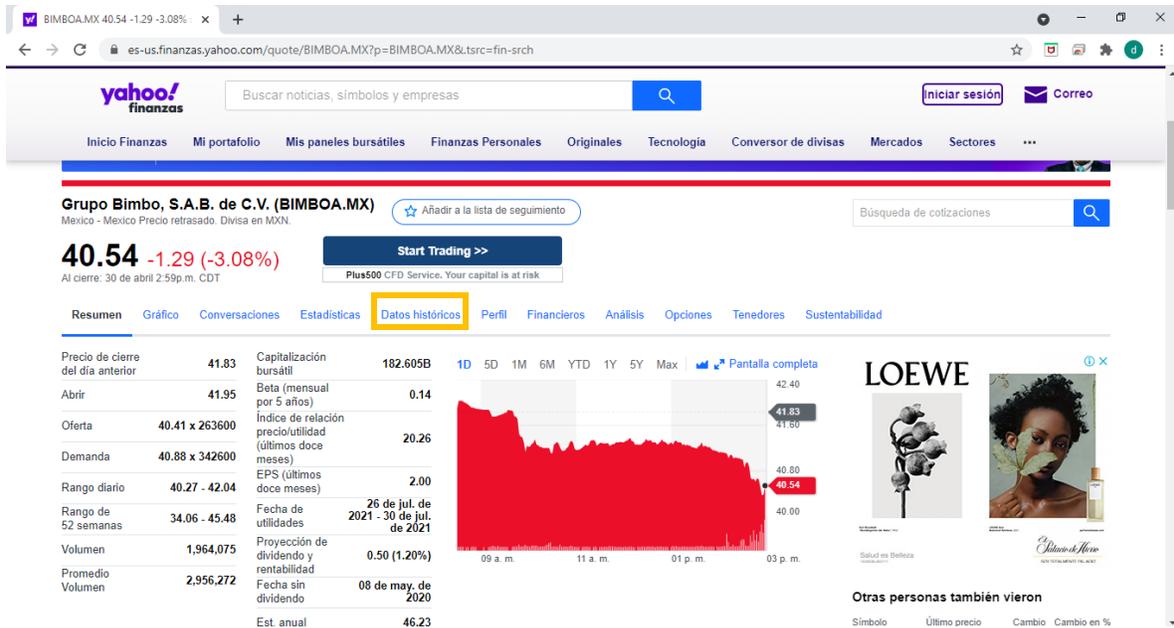


Ilustración 3 Resumen de la empresa

4. Seleccionar el periodo de tiempo y la frecuencia de los datos, en Mostrar, dejar seleccionada la opción Precios Históricos. En este caso el periodo de tiempo se seleccionó desde el primer Lunes del 2019 (7 de Enero de 2019) al Lunes actual (se coloca el límite hasta el 3 de Mayo de 2021 para que se muestren los datos de la semana del 26 de Abril) y frecuencia semanalmente. De esta manera se tendrán datos de un poco más de 2 años. Dar clic en aplicar.

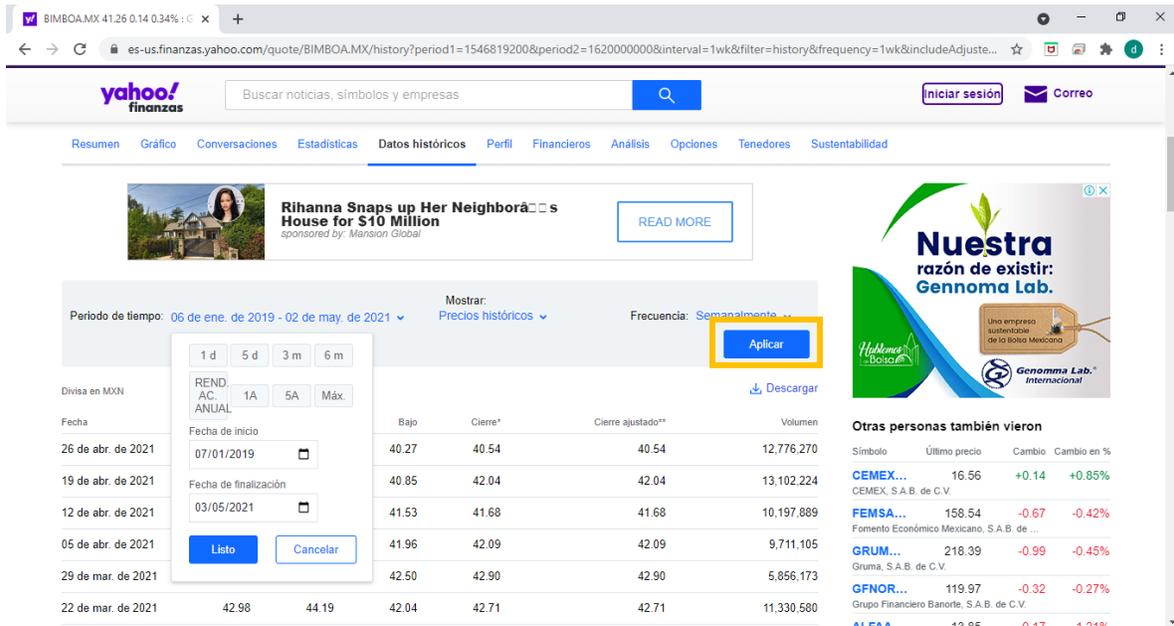


Ilustración 4 Intervalo de los datos

5. Después de visualizar los datos de forma correcta, dar clic en descargar.

Yahoo! finanzas

Buscar noticias, símbolos y empresas

Iniciar sesión Correo

Inicio Finanzas Mi portafolio Mis paneles bursátiles Finanzas Personales Originales Tecnología Conversor de divisas Mercados Sectores

Periodo de tiempo: 06 de ene. de 2019 - 02 de may. de 2021

Mostrar: Precios históricos

Frecuencia: Semanalmente

Aplicar

Divisa en MXN

Fecha Abrir Alto Bajo Cierre\* Cierre ajustado\*\* Volumen

Fecha	Abrir	Alto	Bajo	Cierre*	Cierre ajustado**	Volumen
26 de abr. de 2021	42.10	42.50	40.27	40.54	40.54	12,776.270
19 de abr. de 2021	41.97	43.07	40.85	42.04	42.04	13,102.224
12 de abr. de 2021	41.95	42.57	41.53	41.68	41.68	10,197.889
05 de abr. de 2021	43.33	43.88	41.96	42.09	42.09	9,711.105
29 de mar. de 2021	43.10	44.40	42.50	42.90	42.90	5,856.173
22 de mar. de 2021	42.98	44.19	42.04	42.71	42.71	11,330.580
15 de mar. de 2021	43.03	44.30	42.45	42.92	42.92	21,431.654
08 de mar. de 2021	41.40	43.80	41.03	43.03	43.03	17,019.606
01 de mar. de 2021	39.58	41.70	39.07	41.48	41.48	17,250.714

Descargar

Otras personas también vieron

Símbolo	Último precio	Cambio	Cambio en %
CEMEX...	16.56	+0.14	+0.85%
FEMSA...	158.54	-0.67	-0.42%
GRUM...	218.39	-0.99	-0.45%
GFNOR...	119.97	-0.32	-0.27%
ALFAA...	13.85	-0.17	-1.21%

Ilustración 5 Descargar datos

6. Se descargará un archivo .csv con todos los datos consultados.

Autoguardado BIMBOA.MX - Excel

Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Ayuda

Portapapeles Fuente Alineación Número Estilos

A1	Date	Open	High	Low	Close	Adj Close	Volume
1	07/01/2019	39.189999	39.869999	38.32	38.59	37.638973	10500008
2	14/01/2019	38.630001	39.119999	37.84	38.299999	37.356117	11834479
3	21/01/2019	38.25	38.939999	37.509998	37.900002	36.965977	7977279
4	28/01/2019	37.75	38.380001	36.400002	36.849998	35.941853	12768131
5	04/02/2019	36.849998	37.91	36.73	36.84	35.932098	6413184
6	11/02/2019	36.700001	37.599998	36	36.639999	35.73703	10713247
7	18/02/2019	36.509998	38	35.57	37.669998	36.741646	16779078
8	25/02/2019	37.669998	38.98	37.290001	38.349998	37.404884	26576836
9	04/03/2019	38.439999	39.240002	37.259998	37.959999	37.024498	15534993
10	11/03/2019	38.099998	39.189999	37.84	38.790001	37.834045	26181091
11	18/03/2019	38.790001	39.990002	38.23	39.09	38.126652	17006601
12	25/03/2019	39.099998	40.990002	38.900002	40.580002	39.579933	17750452
13	01/04/2019	40.98	42.43	40.5	42.110001	41.072224	12870906
14	08/04/2019	42.389999	42.389999	39.970001	40.279999	39.287323	10672589
15	15/04/2019	40.240002	41.34	39.599998	40.52	39.521408	6564104
16	22/04/2019	40.52	42.290001	40	42.09	41.052715	12843623

Ilustración 6 Datos descargados

7. Repetir los pasos 2 a 6 para cada serie de datos considerada.

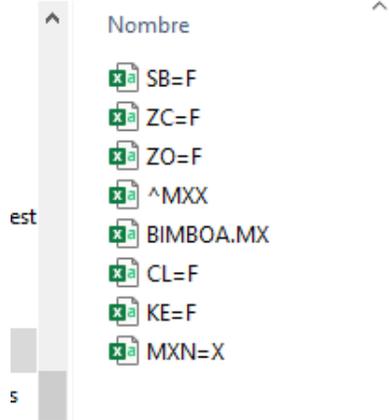


Ilustración 7 Series de datos descargadas

8. Después de descargar las series de datos, en un nuevo archivo .csv, pegar las variables que serán consideradas de cada serie de datos. El nombre de las columnas se asignó de la siguiente manera:

[nombre de la variable].[ nombre de la empresa/serie de datos]

Por ejemplo: Open.BIMBOA.MX

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Date	Open.BIMB	High.BIMC	Low.BIMB	Adj.Close.B	Volume.BIN	High.^MXX	Low.^MXX	High.MXN=	Low.MXN=X	High.CL=F	Low.CL=F	High.ZC=F	Low.ZC=F	High.ZO=F	Low.ZO=F	High.KE=F	Low.KE=F	High.SB=F	Low.SB=F	Close.B
2	07/01/2019	39.189999	39.869999	38.32	37.638973	10500008	44046.77	42313.059	19.468281	19.046301	53.310001	48.110001	384.5	376	295	276.5	511	496.25	12.94	11.94	3
3	14/01/2019	38.630001	39.119999	37.84	37.356117	11834479	44507.398	43171.02	19.200899	18.8652	53.919998	50.380001	383	371	303.5	291.75	509.25	492.25	13.27	12.63	38.296
4	21/01/2019	38.25	38.939999	37.509998	36.965977	7977279	44541.41	43382.352	19.236771	18.8734	54.240002	51.799999	382.25	375.75	299.75	289	518	502.75	13.08	12.37	37.900
5	28/01/2019	37.75	38.380001	36.400002	35.941853	12768131	44226.172	43084.449	19.199006	18.952999	55.66	51.330002	381.5	375.75	294.5	278	513.25	497	12.93	12.46	36.845
6	04/02/2019	36.849998	37.91	36.73	35.932098	6413184	44495.02	43116.559	19.18214	19.008499	55.75	51.330002	381.75	374	294.5	283.5	513.5	492.5	12.93	12.51	3
7	11/02/2019	36.700001	37.599998	36	35.73703	10713247	43616.031	42066.078	19.464399	19.0602	55.869999	51.23	379	371.75	287.75	275	495.5	472	13.25	12.52	36.636
8	18/02/2019	36.509998	38	35.57	36.741646	16779078	44001.219	42803.879	19.299009	19.089999	57.810001	55.290001	378.5	368.5	278	263.25	478.5	449.5	13.5	13.14	37.666
9	25/02/2019	37.669998	38.98	37.290001	37.404884	26576836	44070.129	42291.941	19.3335	19.015799	57.880001	55.02	378.5	357.5	273.75	250.5	462.75	432	13.41	12.61	38.946
10	04/03/2019	38.439999	39.240002	37.259998	37.024498	15534993	42648.699	41140.32	19.599791	19.210581	57.189999	54.52	368	354	259.75	254.25	448.5	419	12.73	12.08	37.956
11	11/03/2019	38.099998	39.189999	37.84	37.834045	26181091	42309.969	41469.922	19.51623	19.1866	58.950001	55.959999	373.75	352.5	277	253.25	444.5	414.25	12.69	12.17	38.790
12	18/03/2019	38.790001	39.990002	38.23	38.126652	17006601	43362.91	42132.301	19.22435	18.7668	60.389999	58.049999	380.75	367.75	284.5	271	452	433.5	12.87	12.43	3
13	25/03/2019	39.099998	40.990002	38.900002	39.579953	17750452	43316.672	42276.039	19.43856	18.9834	60.73	58.169998	380.75	356	280	264.5	454	427.25	12.73	12.33	40.586
14	01/04/2019	40.98	42.43	40.5	41.072224	12870906	45023.539	43168.059	19.374599	19.0366	63.34	60.330001	366.25	356.5	291	271.5	443.5	422.75	12.84	12.29	42.110
15	08/04/2019	42.389999	42.389999	39.970001	39.287323	10672589	45513.898	44373.488	19.123301	18.735001	64.790001	63.130001	363.5	355.25	289.75	276.5	435.25	422	12.83	12.49	40.272
16	15/04/2019	40.240002	41.34	39.599998	39.521408	6564104	45599.57	44347.531	18.995399	18.7363	64.610001	62.990002	364.25	356.75	305	282.25	434.75	414.25	12.79	12.31	4
17	22/04/2019	40.52	42.290001	40	41.052715	12843623	45590.441	44666.25	19.184999	18.7416	66.599998	62.279999	359	342.25	310	289.5	420	398.75	12.8	12.19	4
18	29/04/2019	42.220001	43.330002	40.43	41.833	13127746	45211.148	44187.469	19.156521	18.818199	64.75	60.950001	363.25	349.25	320	297.75	404	383.5	12.39	11.55	42.888
19	06/05/2019	42.98	43.75	40.75	40.808876	11570494	44175.941	42591.93	19.3095	18.937799	62.950001	60.040001	359.25	338.25	302.75	291.5	395.75	379.25	12.15	11.64	4
20	13/05/2019	41.84	42.790001	40.52	40.776794	14754824	43696.328	42524.211	19.250441	18.9709	63.639999	60.639999	384.75	335.5	299.75	270	425.5	378.25	12.06	11.43	41.366
21	20/05/2019	41.349998	41.59	39.099998	39.071602	8901235	43802.25	42299.23	19.157249	18.9794	63.810001	57.330002	404.75	385.75	322.75	292	453	421.25	11.95	11.36	39.636
22	27/05/2019	39.400002	40.5	38.279999	38.982891	19155413	43464.238	42050.422	19.817711	19.003801	59.700001	53.049999	438	407.5	325	306.75	490	445.25	12.13	11.6	39.946
23	03/06/2019	39.580002	41.700001	38.98	40.727512	11523937	43583.988	42748.738	19.8342	19.499411	54.630001	50.599998	436	407	319	290	497	440.75	12.53	11.95	4
24	10/06/2019	41.630001	41.880001	38.860001	39.032177	10617200	43958.398	42971.762	19.297001	19.0874	54.84	50.720001	457.25	408.25	312.75	291.75	478	439	12.78	12.27	39.596
25	17/06/2019	39.740002	41.549999	39.09	39.446156	19389445	43920.789	42908.789	19.20805	18.886499	57.98	51.5	464.25	436.5	307.75	275.25	487	451	12.77	12.17	4
26	24/06/2019	40.169998	41.599998	39.720001	39.456009	8109475	43926.43	43009.852	19.2449	19.0644	59.93	56.75	455	411	288.5	268	478.5	445	12.61	11.98	40.025
27	01/07/2019	40.240002	41.279999	39.810001	39.672855	5084461	43713.43	43019.73	19.20525	18.9517	60.279999	56.040001	435.5	408.25	289.25	275	452	424	12.82	12.25	4
28	08/07/2019	40.110001	40.689999	39.630001	39.731998	6057127	43763.988	42452.949	19.308411	18.8498	60.939999	57.290001	457	428	281	274	461.25	438.25	12.62	12.26	40.310

Ilustración 8 Conjunto de datos construido

## Selección, limpieza y transformación

**Objetivo:** Corregir/eliminar para datos incorrectos, determinación de estrategia para datos incompletos, proyección para solo considerar los atributos relevantes.

## Estrategia para datos incompletos

Como se observa en la ilustración, no se conoce el precio High y Low del dólar en 3 semanas. Se optó por borrar estos registros semanales, ya que no se pueden conseguir los datos faltantes y puede que no sea apropiado borrar toda la columna ya que aún no se ha calculado el coeficiente de correlación y no se sabe si estas variables influyen en el precio de cierre de las acciones de Grupo Bimbo.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
34	19/08/2019	34.369999	34.82	32.799999	32.507099	13468571	40378.73	39352.391	19.92807	19.5944	57.130001	53.240002	370.25	
35	26/08/2019	33.290001	35.599998	32.549999	34.567131	24138979	42680.23	39887.621	20.25243	19.7997	56.889999	52.959999	369	35
36	02/09/2019	35.119999	37.970001	34.009998	36.745441	19126293	43037.262	41724.23	20.1607	19.51	57.759998	52.84	359	3
37	09/09/2019	37.560001	38	35.84	36.203331	11229261	43231.52	42402.648	19.59565	19.467699	58.759998	54	358.5	
38	16/09/2019	36.73	37.540001	35.669998	36.538456	35176318	43665.941	42743.262	null	null	63.380001	57.669998	374.75	36
39	23/09/2019	37.09	38.279999	35.84	35.533081	13342575	43560.051	42563.07	null	null	59.389999	54.75	377.5	36
40	30/09/2019	36.23	36.400002	34.799999	35.266949	7997376	43448.641	42107.012	null	null	56.57	50.990002	392.75	3
41	07/10/2019	35.77	36	34.509998	35.12896	6464756	43496.469	42445.941	19.6311	19.254499	54.93	51.380001	398.75	37
42	14/10/2019	35.580002	35.990002	34	33.571617	8955880	43687.449	43094.031	19.344801	19.0937	54.900002	52.389999	402.5	3
43	21/10/2019	34.200001	35.950001	33.5	35.375374	11309916	43833.391	43203.039	19.184349	19.037399	56.740002	52.709999	393.75	
44	28/10/2019	35.73	37.490002	35.509998	35.887917	11517878	43979.289	43304.859	19.239	19.0068	56.919998	53.709999	391.75	
45	04/11/2019	36.509998	37	35.720001	35.355659	7715307	44261.34	43332.18	19.246401	19.0333	57.880001	55.759998	387.75	3
46	11/11/2019	35.799999	36.66	34.669998	34.675556	6166169	43699.98	43021.301	19.51955	19.0707	57.970001	56.200001	378	37
47	18/11/2019	35.18	35.450001	33.66	33.561756	7189708	43814.5	43101.898	19.495541	19.1702	58.740002	54.759998	373	36
48	25/11/2019	34.099998	34.540001	33	33.315342	23997521	43727.969	42755.988	19.655849	19.3009	58.68	55.02	372.75	36
49	02/12/2019	33.709999	34.09	33	32.566242	6965386	42903.781	41894.75	19.624149	19.286699	59.849998	55.349998	376	36
50	09/12/2019	33	34.599998	32.75	33.26606	7514703	44401.012	41924.211	19.29925	18.979601	60.48	58.110001	374.5	
51	16/12/2019	33.82	34.990002	33.41	34.350285	15051331	45189.828	44118.059	19.079531	18.879999	61.470001	59.709999	390.5	38
52	23/12/2019	34.840001	34.08	33.470001	34.379856	3337688	44590.93	44057.488	18.97851	18.8008	61.970001	60.090002	391	3

Ilustración 9 Manejo de datos perdidos

## Estrategia para considerar los atributos relevantes

### 1. Leer conjunto de datos

```

RGui (64-bit)
Archivo Editar Paquetes Ventanas Ayuda

R Console
> library(MASS)
> library(neuralnet)
> library(e1071)
>
> #para limpiar memoria
> rm(list=ls())
>
> #inicializar generador de numeros aleatorios
> set.seed(42)
>
> #lectura de datos
> datos <- read.csv("E:/8 Semestre/Mineria/Parcial 3/Bimbo/conjuntoDeDatos_BIMBOA.MX.csv",header=TRUE)
> attach(datos)
The following objects are masked from datos (pos = 3):
  Adj.Close.BIMBOA.MX, Close.BIMBOA.MX, Date, High.MXX, High.BIMBOA.MX, High.CL.F, High.KE.F, High.MXN.X, High.SB.F, High.ZC.F, High.ZO.F,
  Low.MXX, Low.BIMBOA.MX, Low.CL.F, Low.KE.F, Low.MXN.X, Low.SB.F, Low.ZC.F, Low.ZO.F, Open.BIMBOA.MX, Volume.BIMBOA.MX

>
> datos[1,]
  Date Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX Volume.BIMBOA.MX High.MXX Low.MXX High.MXN.X Low.MXN.X High.CL.F Low.CL.F High.ZC.F
1 07/01/2019      39.19      39.87      38.32      37.63897      10500008      44046.77 42313.06 19.46828 19.0463 53.31 48.11 384.5
  Low.ZC.F High.ZO.F Low.ZO.F High.KE.F Low.KE.F High.SB.F Low.SB.F Close.BIMBOA.MX
1      376      295      276.5      511      496.25      12.94      11.94      39.59
  
```

Ilustración 10 Leer conjunto de datos

### 2. Calcular el coeficiente de correlación.

```

> #correlación
> datos_cor <- cor(datos)
> datos_cor

```

	Open.BIMBOA.MX	High.BIMBOA.MX	Low.BIMBOA.MX	Adj.Close.BIMBOA.MX	Volume.BIMBOA.MX	High.MXN	Low.MXN	High.MXN.X	Low.MXN.X	High.CL.F
Open.BIMBOA.MX	1.0000000	0.9522028	0.9746843	0.9394654	-0.2371066	0.3290665	0.3734463	-0.2384348	-0.141395539	0.30271785
High.BIMBOA.MX	0.9522028	1.0000000	0.9587948	0.96643313	-0.12881266	0.2210776	0.2486620	-0.08602035	-0.003842232	0.18057449
Low.BIMBOA.MX	0.9746843	0.9587948	1.0000000	0.97212786	-0.27583911	0.3495955	0.3947440	-0.22875128	-0.133871537	0.30833182
Adj.Close.BIMBOA.MX	0.9394654	0.96643313	0.9721279	1.0000000	-0.19365419	0.2736354	0.3086582	-0.116022173	-0.030085449	0.22767467
Volume.BIMBOA.MX	-0.2371107	-0.128812659	-0.2758391	-0.19365419	1.0000000	-0.2160071	-0.2815419	0.257504590	0.210851786	-0.17888753
High.MXN	0.3290665	0.221077559	0.3495955	0.27363539	-0.21600708	1.0000000	0.9797962	-0.771788253	-0.771860950	0.85731385
Low.MXN	0.3734463	0.248661998	0.3947440	0.30865825	-0.28154194	0.9797962	1.0000000	-0.805050976	-0.781748385	0.88294546
High.MXN.X	-0.23843455	-0.08602035	-0.2287571	-0.11602217	0.25750459	-0.7717893	-0.8050510	1.000000000	0.975986623	-0.86181055
Low.MXN.X	-0.14139553	-0.003842232	-0.1338715	-0.03008545	0.21085179	-0.7718610	-0.7817484	0.975986623	1.000000000	-0.84635756
High.CL.F	0.30271785	0.18057449	0.3083318	0.22767467	-0.17888753	0.8573138	0.8829455	-0.861810554	-0.846357563	1.000000000
Low.CL.F	0.3411878	0.24493060	0.3506363	0.27898728	-0.17458517	0.7842562	0.8080043	-0.824076272	-0.804951265	0.53504678
High.ZC.F	0.4870611	0.439120090	0.4807066	0.46654628	-0.11977973	0.6733742	0.6438430	-0.283954682	-0.282693100	0.49503808
Low.ZC.F	0.4673920	0.419267975	0.4685741	0.45434462	-0.12504802	0.7062581	0.6771435	-0.312353706	-0.311225037	0.52227538
High.ZO.F	0.3832538	0.352537772	0.3956146	0.40958507	-0.12470266	0.4922910	0.4648171	0.007476073	0.046175248	0.22944112
Low.ZO.F	0.3899114	0.348261890	0.4038162	0.40269552	-0.18305342	0.5265995	0.5052339	-0.029134740	0.003284635	0.25876146
High.KE.F	0.4463396	0.45390677	0.4544791	0.47190081	-0.09740415	0.3521308	0.3062381	0.077654526	0.07824670	0.03827095
Low.KE.F	0.4443567	0.449040228	0.4585285	0.47170664	-0.10851773	0.3819562	0.3389052	0.046036625	0.052047353	0.06559642
High.SB.F	0.4225262	0.396931478	0.4346192	0.42761051	-0.12246088	0.6422137	0.6017898	-0.314155406	-0.314229050	0.40441733
Low.SB.F	0.3590893	0.342349897	0.3689562	0.35460139	-0.07967494	0.6128054	0.5791712	-0.443277578	-0.428987206	0.46342591
Close.BIMBOA.MX	0.9405995	0.962661640	0.9755956	0.99587431	-0.20089536	0.3022599	0.3400972	-0.165519474	-0.081409273	0.26972514

	Low.CL.F	High.ZC.F	Low.ZC.F	High.ZO.F	Low.ZO.F	High.KE.F	Low.KE.F	High.SB.F	Low.SB.F	Close.BIMBOA.MX
Open.BIMBOA.MX	0.34118779	0.4870611	0.4673920	0.383253765	0.389311430	0.44633957	0.44435670	0.4225262	0.35908930	0.94059951
High.BIMBOA.MX	0.24493060	0.4391201	0.4192680	0.352537772	0.348261890	0.45390677	0.44904023	0.3969315	0.34234989	0.96266164
Low.BIMBOA.MX	0.35063625	0.4807066	0.4685741	0.395614626	0.403816235	0.45447808	0.45852853	0.4346192	0.36895618	0.97559564
Adj.Close.BIMBOA.MX	0.27898728	0.4665463	0.4543446	0.409585072	0.402695522	0.47190081	0.47170664	0.4276109	0.35460139	0.99587431
Volume.BIMBOA.MX	-0.17458517	-0.1197797	-0.1250480	-0.124702660	-0.183053416	-0.09740415	-0.10851773	-0.1224609	-0.07967494	-0.20089536
High.MXN	0.7842562	0.6733742	0.7062581	0.492291025	0.526599454	0.35213078	0.38135622	0.6422137	0.61280538	0.30225995
Low.MXN	0.8080043	0.6438430	0.6771435	0.464817104	0.505233947	0.30623808	0.33890522	0.6017898	0.57917122	0.34009716
High.MXN.X	-0.82407627	-0.2839547	-0.3123537	-0.007476073	-0.029134740	0.07765453	0.04603663	-0.3141554	-0.44327758	-0.16551947
Low.MXN.X	-0.80495126	-0.2826931	-0.3112250	0.046175248	0.003284635	0.0782467	0.05204735	-0.3142290	-0.42898721	-0.08140927
High.CL.F	0.93504678	0.4950831	0.5222754	0.229414124	0.258761458	0.03827095	0.06559842	0.4044173	0.46342591	0.26972514
Low.CL.F	1.00000000	0.4321861	0.4604254	0.187785390	0.219687390	0.02023871	0.04982734	0.3995503	0.42989039	0.31481934
High.ZC.F	0.43218611	1.0000000	0.9900465	0.720980126	0.742441906	0.78937011	0.78630112	0.7893477	0.60543097	0.44805950
Low.ZC.F	0.46042542	0.9900465	1.0000000	0.727028520	0.753457880	0.80460957	0.80474499	0.8201023	0.63891038	0.43674868
High.ZO.F	0.18778539	0.7209801	0.7270285	1.000000000	0.96054428	0.71150393	0.71829871	0.6307971	0.3330064	0.37361032
Low.ZO.F	0.21968739	0.7424415	0.7534579	0.96054428	1.000000000	0.72595818	0.73517093	0.6694912	0.39901259	0.36716034
High.KE.F	0.02023871	0.7953701	0.8046096	0.711502810	0.725958183	1.00000000	0.98941510	0.8064707	0.54802218	0.43289685
Low.KE.F	0.04982734	0.7863011	0.8047450	0.716298711	0.735170927	0.98941510	1.00000000	0.8259337	0.56590669	0.43432281
High.SB.F	0.39955028	0.7893477	0.8201023	0.630797094	0.669491244	0.80647075	0.82593374	1.00000000	0.76684115	0.39989501
Low.SB.F	0.42989039	0.6054310	0.6389104	0.383300637	0.399012587	0.54802218	0.56590669	0.7668412	1.00000000	0.34715604
Close.BIMBOA.MX	0.31481934	0.4480595	0.4367487	0.373610622	0.367160338	0.43289685	0.43432281	0.3998950	0.34715604	1.00000000

Ilustración 11 Coeficiente de correlación

### Coefficientes de correlación con respecto a Close.BIMBOA.MX

	Open.BIMBOA.MX	High.BIMBOA.MX	Low.BIMBOA.MX	Adj.Close.BIMBOA.MX	Volume.BIMBOA.MX
Close.BIMBOA.MX	0.9405995	0.962661640	0.9755956	0.99587431	-0.20089536

Tabla 1 Coeficientes de correlación

	High.MXN	Low.MXN	High.MXN.X	Low.MXN.X	High.CL=F
Close.BIMBOA.MX	0.3022599	0.3400972	-0.165519474	-0.081409273	0.26972514

Tabla 2 Coeficientes de correlación

	Low.CL=F	High.ZC=F	Low.ZC=F	High.ZO=F	Low.ZO=F
Close.BIMBOA.MX	0.31481934	0.4480595	0.4367487	0.373610622	0.367160338

Tabla 3 Coeficientes de correlación

	High.KE=F	Low.KE=F	High.SB=F	Low.SB=F
Close.BIMBOA.MX	0.43289685	0.43432281	0.3998950	0.34715604

Tabla 4 Coeficientes de correlación

3. Crear un nuevo conjunto de datos con las columnas que tienen un coeficiente de correlación mayor a | 0.3 | con la variable Close.BIMBOA.MX

```

> #quitar columnas con baja correlación
> datos2 <- datos[, -5]
> datos2 <- datos2[, -7]
> datos2 <- datos2[, -7]
> datos2 <- datos2[, -7]
> datos2[1,]
Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX High..MX
1 39.19 39.87 38.32 37.63897 44046.77
Low..MX Low.CL.F High.ZC.F Low.ZC.F High.ZO.F Low.ZO.F High.KE.F Low.KE.F
1 42313.06 48.11 384.5 376 295 276.5 511 496.25
High.SB.F Low.SB.F Close.BIMBOA.MX
1 12.94 11.94 38.59
> |

```

Ilustración 12 Crear un nuevo conjunto de datos

## Minería de datos

**Objetivo:** Elección de tarea a realizar y el método a utilizar.

### Tarea de minería de datos

La tarea de minería de datos de este proyecto es predicción, ya que a partir de un conjunto de datos históricos se pronosticarán valores futuros. En este caso se pronosticará el precio de cierre de las acciones de Grupo Bimbo.

### Tipo de modelo

El modelo considerado en este proyecto es un modelo predictivo, específicamente será una red neuronal.

### Algoritmo

El algoritmo utilizado en la red neuronal es el Algoritmo Backpropagation, de esta manera se entrenará la red neuronal con aprendizaje supervisado con una función real que asigna a cada instancia un valor real, con el objetivo de minimizar el error entre el valor pronosticado y el valor real.

### Ejecución de código en R

Se desarrollaron 3 modelos con diferente número de neuronas en capa oculta para encontrar el mejor modelo (el de menor error y mayor eficiencia) y con este pronosticar el precio de cierre de las acciones de Grupo Bimbo.

#### 1. Normalizar el conjunto de datos.

```

> #normalización
> datos_nor <- scale(datos2)
> datos_nor
Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX High..MX Low..MX Low.CL.F High.ZC.F Low.ZC.F High.ZO.F Low.ZO.F
[1,] 0.287075693 0.10321460 0.339544276 -0.02576457 0.49449086 0.354723077 0.07618097 -0.32681296 -0.19436240 -0.53200690 -0.566222105
[2,] 0.134933001 -0.10415254 0.211520128 -0.10170337 0.56463696 0.581932466 0.23814991 -0.34560935 -0.26496349 -0.28318297 -0.105017317
[3,] 0.031692677 -0.15392065 0.123502993 -0.20644466 0.57424663 0.637898443 0.33821845 -0.35500754 -0.19789246 -0.39295823 -0.188185393
[4,] -0.104149497 -0.30875423 -0.172551783 -0.48139303 0.48517920 0.559006496 0.30509734 -0.36440574 -0.19789246 -0.54664360 -0.520857700
[5,] -0.348665954 -0.43870458 -0.084535715 -0.48401197 0.56113942 0.567509855 0.33821845 -0.36127301 -0.22260284 -0.54664360 -0.354521547
[6,] -0.389417792 -0.52441688 -0.279239107 -0.53638218 0.31279005 0.289316510 0.29805011 -0.39573306 -0.25437333 -0.74423907 -0.611586510
[7,] -0.441038633 -0.41382052 -0.393927406 -0.26667136 0.42162094 0.484704502 0.58416204 -0.40199852 -0.30026403 -1.02965475 -0.966941019
[8,] -0.125884789 -0.14286080 0.064826058 -0.08861081 0.44109082 0.349130723 0.56513483 -0.40199852 -0.45558643 -1.15406672 -1.352538465
[9,] 0.083312431 -0.07097297 0.056823749 -0.19073364 0.03948004 0.044152821 0.52989937 -0.53357326 -0.50500719 -1.56389437 -1.239127451
[10,] -0.009060519 -0.08479827 0.211520128 0.02660671 -0.05622488 0.131439512 0.63137740 -0.46152043 -0.52618751 -1.05892816 -1.269370388
[11,] 0.178402497 0.13639417 0.315539748 0.10516337 0.24127322 0.306853915 0.77866159 -0.37380394 -0.31085420 -0.83937763 -0.732558258
[12,] 0.262623830 0.41288369 0.494240655 0.49532799 0.22820905 0.344919385 0.78711803 -0.37380394 -0.47676675 -0.97110795 -0.929137348
[13,] 0.773390948 0.81102804 0.920987282 0.89596568 0.71046734 0.581148336 0.92524120 -0.55550238 -0.46970664 -0.64910051 -0.717436789
[14,] 1.156465608 0.79996818 0.779627552 0.41677052 0.84901350 0.900376090 1.13665391 -0.58996243 -0.48735692 -0.68569226 -0.566222105
[15,] 0.572345074 0.50965446 0.680941471 0.47961569 0.87321923 0.893502023 1.12678805 -0.58056424 -0.46617659 -0.23927266 -0.392325218

```

Ilustración 13 Normalización

2. Separar el conjunto de datos en un conjunto de entrenamiento y prueba. En el conjunto de entrenamiento cuenta el 70% de los datos y el de prueba el 30%.

```
>
> #separar conjunto de entrenamiento y prueba
> tamaño_muestra <- (trunc(0.7*nrow(datos_nor)))
> muestra <- sample(1:nrow(datos_nor),tamaño_muestra)
> entrena <- datos_nor[muestra,]
> prueba <- datos_nor[-muestra,]
> entrena
```

*Ilustración 14 Separar conjunto de datos*

3. Creación del Modelo 1, red neuronal con 5 neuronas. Se calcula el valor estimado de la variable Close.BIMBOA.MX y el error con respecto a los valores reales de esta variable.

```
.
> # ----- MODELO 1 -----
> #red de 5 neuronas
> red5 <- neuralnet(Close.BIMBOA.MX~Open.BIMBOA.MX+High.BIMBOA.MX+Low.BIMBOA.MX+$
> plot(red5)
> plot(red5)
>
> valor_estimado_Modelo1 <- (compute(red5, prueba))$net.result
>
> #regresar a unidades originales
> media <- mean (datos2[,16])
> desviacion <- sd (datos2[,16])
>
> valor_estimado_en_mxn_Modelo1 <- (valor_estimado_Modelo1*desviacion)+media
>
> #error
>
> diferencias <- datos2[-muestra,16]-valor_estimado_en_mxn_Modelo1
>
> dif_cuad <- diferencias^2
>
> ECM_Modelo1 <- mean(dif_cuad)
> ECM_Modelo1
[1] 0.2140976
> |
```

*Ilustración 15 Modelo 1, 5 neuronas*

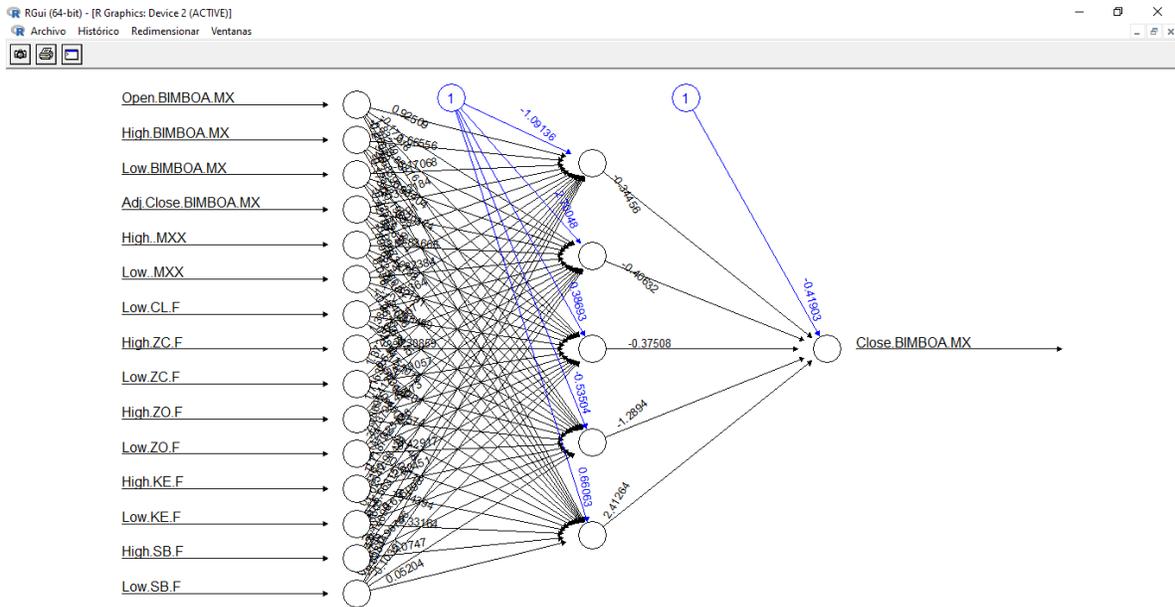


Ilustración 16 Modelo 1, 5 neuronas

4. Creación del Modelo 2, red neuronal con 10 neuronas. Se calcula el valor estimado de la variable Close.BIMBOA.MX y el error con respecto a los valores reales de esta variable.

```

>
> # ----- MODELO 2 -----
>
> #red de 10 neuronas
> red10 <- neuralnet(Close.BIMBOA.MX~Open.BIMBOA.MX+High.BIMBOA.MX+Low.BIMBOA.MX$
> plot(red10)
> valor_estimado_Modelo2 <- (compute(red10, prueba))$net.result
>
> #regresar a unidades originales
> valor_estimado_en_mxn_Modelo2 <- (valor_estimado_Modelo2*desviacion)+media
>
> #error
>
> diferencias <- datos2[-muestra,16]-valor_estimado_en_mxn_Modelo2
>
> dif_cuad <- diferencias^2
>
> ECM_Modelo2 <- mean(dif_cuad)
>
> ECM_Modelo2
[1] 0.4756798
>

```

Ilustración 17 Modelo 2, 10 neuronas

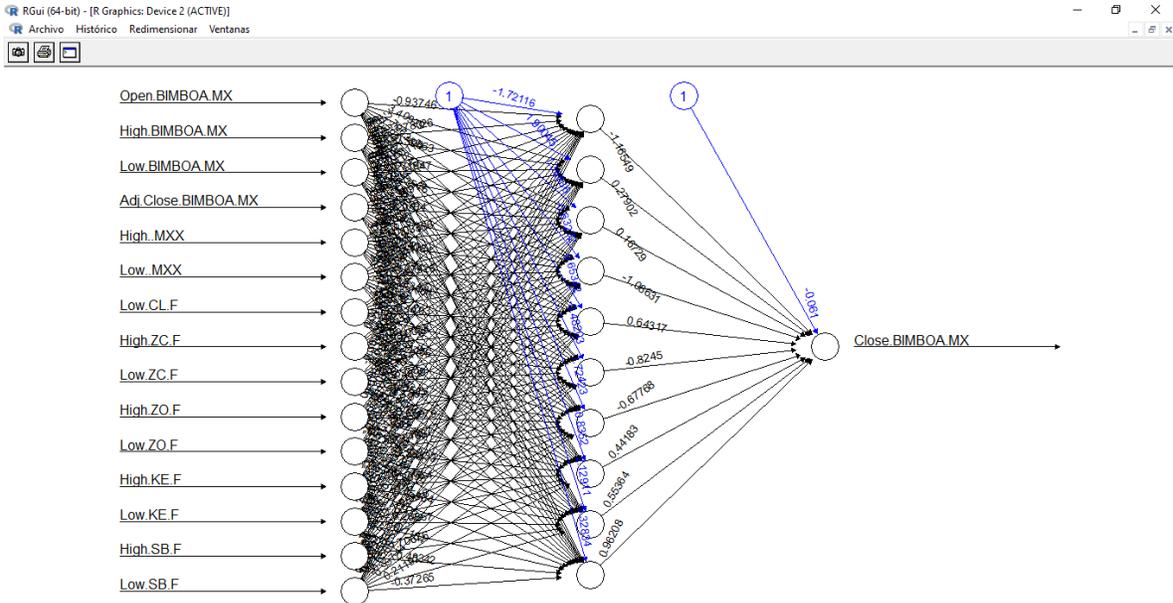


Ilustración 18 Modelo 2, 10 neuronas

5. Creación del Modelo 3, red neuronal con 15 neuronas. Se calcula el valor estimado de la variable Close.BIMBOA.MX y el error con respecto a los valores reales de esta variable.

```

>
> # ----- MODELO 3 -----
>
> #red de 15 neuronas
> red15 <- neuralnet(Close.BIMBOA.MX~Open.BIMBOA.MX+High.BIMBOA.MX+Low.BIMBOA.MX$
> plot(red15)
Error in UseMethod("depth") :
  no applicable method for 'depth' applied to an object of class "NULL"
>
> valor_estimado_Modelo3 <- (compute(red15, prueba))$net.result
>
> #regresar a unidades originales
> valor_estimado_en_mxn_Modelo3 <- (valor_estimado_Modelo3*desviacion)+media
>
> #error
>
> diferencias <- datos2[-muestra,16]-valor_estimado_en_mxn_Modelo3
>
> dif_cuad <- diferencias^2
>
> ECM_Modelo3 <- mean(dif_cuad)
>
> ECM_Modelo3
[1] 0.7280751
>

```

Ilustración 19 Modelo 3, 15 neuronas

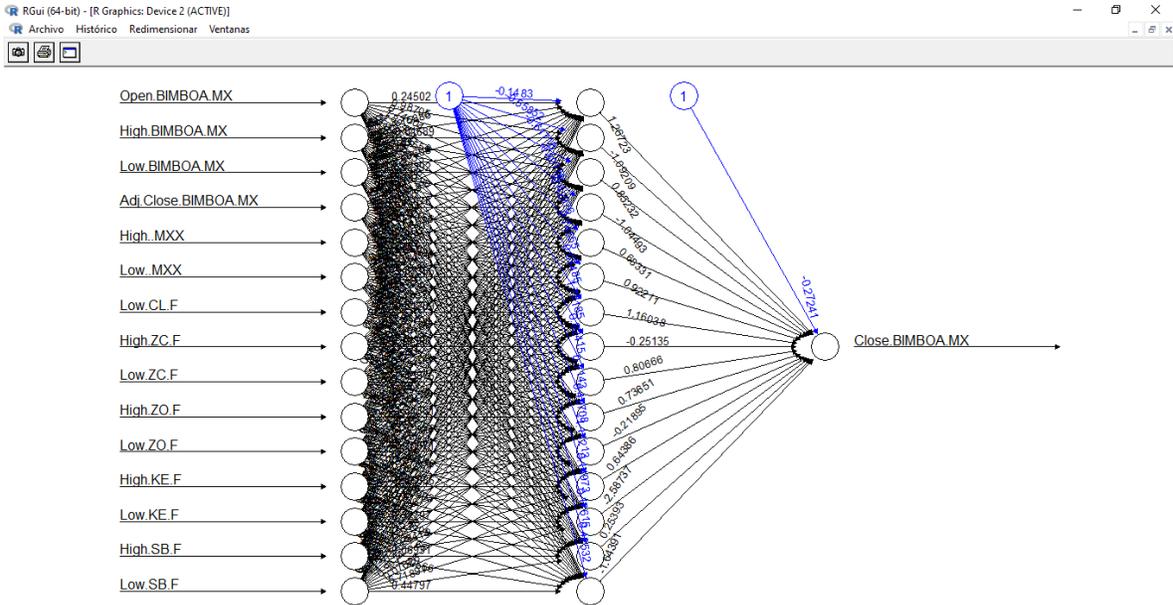


Ilustración 20 Modelo 3, 15 neuronas

## Evaluación e interpretación

**Objetivo:** Evaluación y análisis de los modelos creados y patrones identificados por expertos.

### Técnicas de evaluación

Como se observó en la fase anterior, para entrenar y probar un modelo, los datos se partieron en dos conjuntos: conjunto de entrenamiento y conjunto de prueba, de esta manera se garantiza que la validación de la precisión del modelo es una medida independiente.

### Medidas de evaluación

Para evaluar los modelos creados se utilizará el error cuadrado medio, para posteriormente comparar estos valores y seleccionar el modelo con el menor error cuadrado medio, ya que este será el mejor modelo.

### Modelo final construido

Para elegir el mejor modelo se tomó en cuenta el error cuadrado medio de cada modelo, se elige el modelo con el menor error. En este caso el mejor modelo es el Modelo 1, el cual tiene 5 neuronas en capa oculta.

Con este modelo se calcula el porcentaje de eficiencia, tomando en cuenta los valores estimados por el modelo y los valores reales del conjunto de prueba.

```

>
> #errores
> ECM_Modelo1
[1] 0.2140976
> ECM_Modelo2
[1] 0.4756798
> ECM_Modelo3
[1] 0.7280751
>
> #en porcentajes
> #usar modelo con menor error
> diferencias <- datos2[-muestra,16]-valor_estimado_en_mxn_Modelo1
>
> porcentajes <-abs(diferencias)*100/datos2[-muestra,16]
>
> promedio_porcentajes <- mean(porcentajes)
>
> eficiencia <- 100-promedio_porcentajes
>
> eficiencia
[1] 98.96433
> |

```

Ilustración 21 Porcentaje de eficiencia del mejor modelo

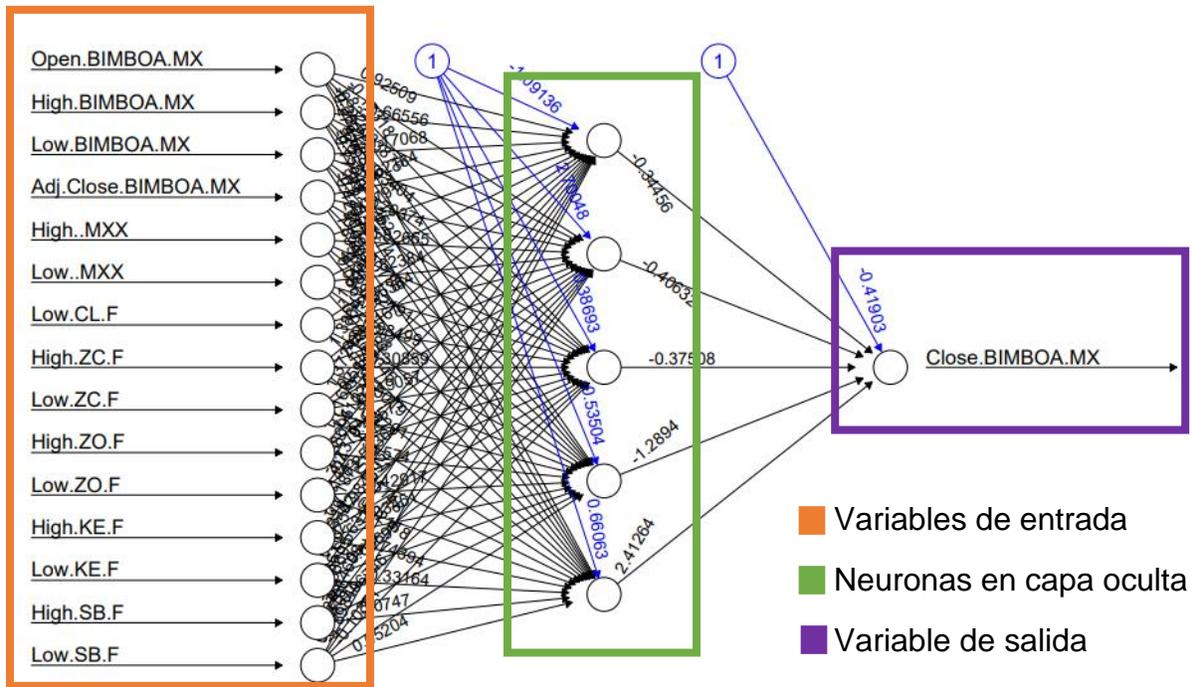


Ilustración 22 Red neuronal propuesta

### Variables de entrada

Las variables de entrada fueron seleccionadas tomando en cuenta el índice de correlación, estas variables tienen una correlación mayor a  $| 0.3 |$  (valor absoluto), ya que esto indica que tienen una correlación débil a fuerte.

- Open.BIMBOA.MX

- High.BIMBOA.MX
- Low.BIMBOA.MX
- Adj.Close.BIMBOA.MX
- High..MXX
- Low..MXX
- Low.CL.F
- High.ZC.F
- Low.ZC.F
- High.ZO.F
- Low.ZO.F
- High.KE.F
- Low.KE.F
- High.SB.F
- Low.SB.F

### **Número de neuronas en capa oculta**

El número de neuronas en capa oculta de la red neuronal propuesta es 5, ya que de los tres modelos construidos, este modelo tuvo el menor error cuadrado medio y por ende tuvo el mejor porcentaje de eficiencia.

### **Variable de salida**

Como se mencionó anteriormente, la variable que se pronosticará es el precio al cierre de las acciones de Grupo Bimbo (Close.BIMBOA.MX), ya que de esa manera se está pronosticando cómo va a cerrar la acción y se tiene todo el día financiero para tomar decisiones acerca de comprar o vender acciones.

### **Interpretación y contextualización**

#### **Información del modelo**

Error cuadrado medio: 0.2140976

Porcentaje de eficiencia: 98.96433 %

Como se puede observar, el porcentaje de eficiencia del modelo construido es bastante alto, pero debido al contexto de este proyecto (precios de las acciones de la bolsa mexicana de valores), puede haber variación entre los precios pronosticados y los precios reales de las acciones.

Esto ocurre ya que no se tienen muchos datos específicos de Grupo Bimbo, como el número de ventas, ganancias, número de empleados, número de puntos de venta, cambios organizacionales, etc. (Todos estos datos semanales, ya que este fue el intervalo de las variables de entrada de la red). Este tipo de variables pueden influir en el precio de las acciones, pero al no tener acceso a estas, puede que al pronosticar el precio de cierre de las acciones haya una variación entre el precio que se obtenga al ejecutar la red y el precio real en la bolsa de valores.

Debido a este contexto, el precio pronosticado por esta red neuronal solo se tomará como referencia, así que se monitoreará el precio de las acciones durante todo el día financiero y se tomarán las decisiones de compra de la siguiente manera:

1. Si el precio de las acciones no baja durante el día, se esperará a realizar la compra aproximadamente a las 2:30 p.m., ya que será cercana a la hora de cierre y se tiene una referencia del precio de cierre.
2. Si durante el día financiero el precio de las acciones baja y se acerca por lo menos un poco al precio pronosticado, se realizará la compra.

Estas decisiones se tomarán con la información y gráficos que muestra el portal Yahoo Finanzas.

### Difusión, uso y monitorización

En esta fase se aplicó el modelo a diferentes conjuntos de datos. Durante la semana, se obtuvieron los datos de las variables de entrada de la semana actual (datos acumulados a un día antes de la ejecución del modelo) para colocarlos como entrada en la red neuronal y obtener el precio de cierre de las acciones.

La red neuronal se ejecutó aproximadamente al inicio del día financiero para monitorear el precio de las acciones durante el día y realizar las decisiones de compra.

### 6 de Mayo de 2021

1. Se obtuvieron los datos acumulados de la semana actual (3 de Mayo de 2021) al 5 de Mayo de 2021 de las variables de entrada.

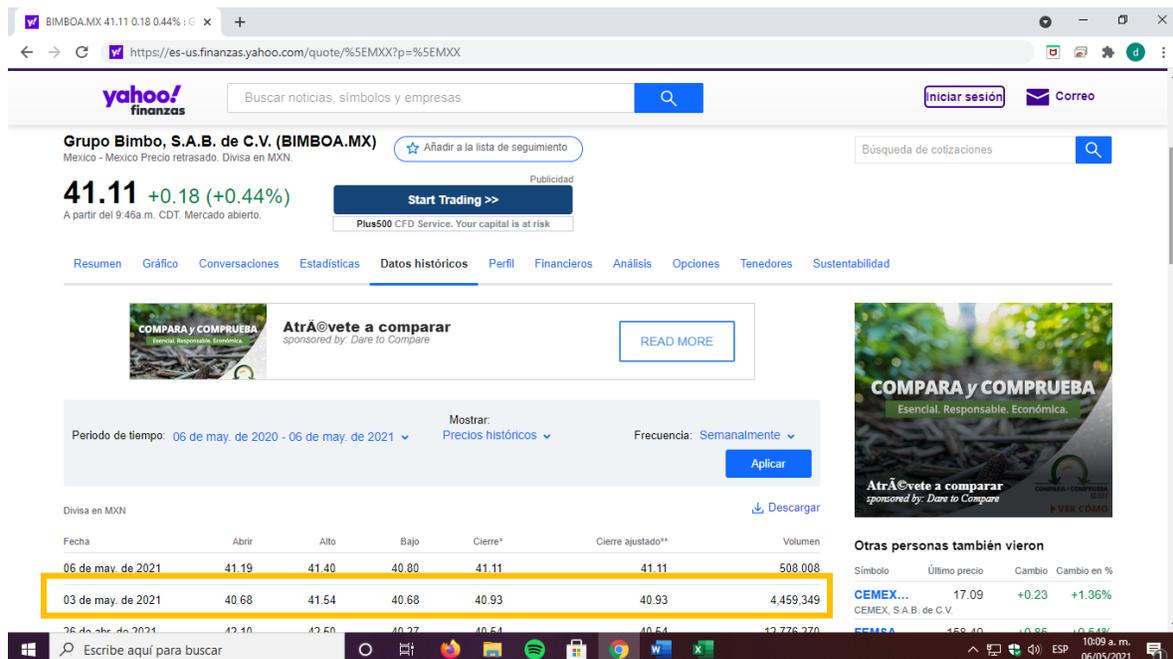
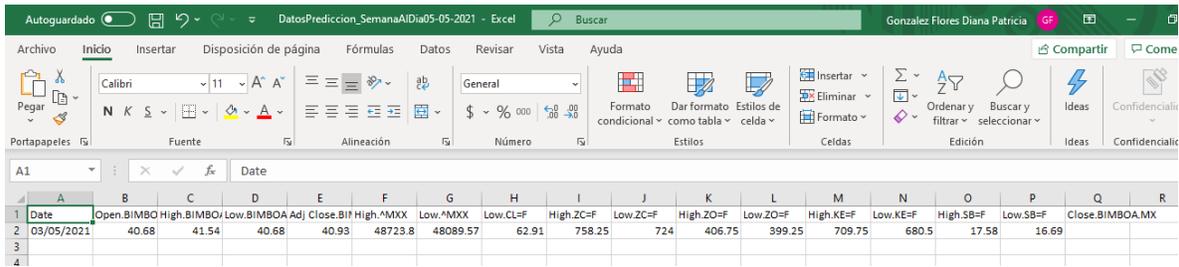


Ilustración 23 Datos semanales de Grupo Bimbo acumulados al 5 de Mayo de 2021

## 2. Estos datos se colocaron en un archivo .csv.



Date	Open.BIMBOA.MX	High.BIMBOA.MX	Low.BIMBOA.MX	Adj.Close.BIMBOA.MX	Close.BIMBOA.MX	High..MXX	Low..MXX	Low.CL.F	High.ZC.F	Low.ZC.F	High.ZO.F	Low.ZO.F	High.KE.F	Low.KE.F	High.SB.F	Low.SB.F	Close.BIMBOA.MX
03/05/2021	40.68	41.54	40.68	40.93	48723.8	48089.57	62.91	758.25	724	406.75	399.25	709.75	680.5	17.58	16.69		

Ilustración 24 Datos de entrada para ejecución de red neuronal

3. En R estos datos se ingresaron como entrada en la red neuronal y se pronosticó el precio de cierre correspondiente a estos datos. Esto se realiza de la misma manera en que se ejecutó la red neuronal con el conjunto de datos de prueba.

```
> #predecir valores nuevos
> #lectura de datos
> datosPred <- read.csv("E:/8 Semestre/Mineria/Final/DatosPrediccion_SemanaAlDi$
> attach(datosPred)
The following objects are masked from datos:

  Adj.Close.BIMBOA.MX, Close.BIMBOA.MX, Date, High..MXX,
  High.BIMBOA.MX, High.KE.F, High.SB.F, High.ZC.F, High.ZO.F,
  Low..MXX, Low.BIMBOA.MX, Low.CL.F, Low.KE.F, Low.SB.F, Low.ZC.F,
  Low.ZO.F, Open.BIMBOA.MX

>
> datosPred[1,]
  Date Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX
1 03/05/2021          40.68          41.54          40.68          40.93
  High..MXX Low..MXX Low.CL.F High.ZC.F Low.ZC.F High.ZO.F Low.ZO.F High.KE.F
1 48723.8 48089.57 62.91 758.25 724 406.75 399.25 709.75
  Low.KE.F High.SB.F Low.SB.F Close.BIMBOA.MX
1 680.5 17.58 16.69 NA
>
> #quitar columna de fechas
> datosPred <- datosPred[,-1]
```

Ilustración 25 Pronóstico de precio de cierre en R

```

<
> #quitar columna de fechas
> datosPred <-datosPred[,-1]
> datosPred[1,]
Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX High..MX
1 40.68 41.54 40.68 40.93 48723.8
Low..MX Low.CL.F High.ZC.F Low.ZC.F High.ZO.F Low.ZO.F High.KE.F Low.KE.F
1 48089.57 62.91 758.25 724 406.75 399.25 709.75 680.5
High.SB.F Low.SB.F Close.BIMBOA.MX
1 17.58 16.69 NA
>
> #normalización
> media_original <- attr(datos_nor,"scaled:center")
> desviacion_original <- attr(datos_nor,"scaled:scale")
> datos_pred_nor <- (datosPred-media_original)/desviacion_original
> datos_pred_nor
Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX High..MX
1 0.6918856 0.5649524 0.9689963 0.8577826 1.755939
Low..MX Low.CL.F High.ZC.F Low.ZC.F High.ZO.F Low.ZO.F High.KE.F Low.KE.F
1 1.884487 1.12115 4.356621 4.719473 2.739296 3.146098 2.856541 3.069269
High.SB.F Low.SB.F Close.BIMBOA.MX
1 2.217405 1.570323 NA
>
>
> valor_estimado_ModeloFinal <- (compute(red5, datos_pred_nor))$net.result
>
> #regresar a unidades originales
> valor_estimado_en_mxn_ModeloFinal <- (valor_estimado_ModeloFinal*desviacion)+$
> valor_estimado_en_mxn_ModeloFinal
[1,]
[1,] 39.35061
> |
<

```

Ilustración 26 Pronóstico de precio de cierre en R

### Precio de cierre pronosticado: 39.35061 MXN

Con este valor se monitoreó el precio de las acciones de Grupo Bimbo durante el día financiero. Como se muestra en la imagen, el precio se mantuvo estable durante el día, por lo que se decidió hacer la compra más cercana a la hora del cierre de la bolsa.

**Grupo Bimbo, S.A.B. de C.V. (BIMBOA.MX)**  
Mexico - Mexico Precio retrasado. Divisa en MXN.

★ Añadir a la lista de seguimiento

**41.10** +0.17 (+0.42%)  
A partir del 10:21a.m. CDT. Mercado abierto.

Start Trading >>

Plus500 CFD Service. Your capital is at risk

Resumen Gráfico Conversaciones Estadísticas Datos históricos Perfil Financieros Análisis Opciones Tenedores Sust

Precio de cierre del día anterior	40.93	Capitalización bursátil	185.128B
Abrir	41.19	Beta (mensual por 5 años)	0.15
Oferta	41.09 x 263600	Índice de relación precio/utilidad (últimos doce meses)	15.05
Demanda	41.10 x 342600	EPS (últimos doce meses)	N/D
Rango diario	40.80 - 41.40	Fecha de utilidades	N/D
Rango de 52 semanas	34.22 - 45.48	Proyección de dividendo y rentabilidad	1.00 (2.44%)
Volumen	725,100	Fecha sin dividendo	13 de may. de 2021
Promedio Volumen	2,889,911	Est. anual	N/D

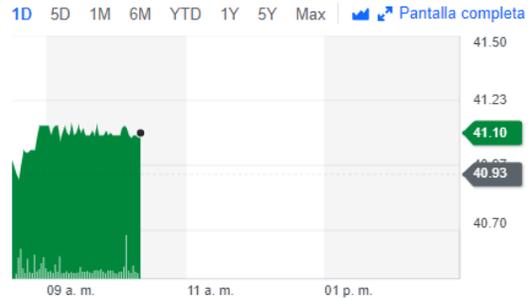


Ilustración 27 Monitoreo de precio de acciones

A las 2:50 p.m. se realizó la compra de 25 títulos (acciones)

Ilustración 28 Compra de acciones 6 de Mayo de 2021

Operación

ACCIONES

FONDOS DE INVERSIÓN

MOVIMIENTOS DEL DÍA

HISTÓRICO DE MOVIMIENTOS

Bienvenido Diana Patricia | Último Log In: May 6, 2021, 12:01 PM | Cambio de Password

Alta de Orden Exitosa

Confirmación

Su Orden fue correctamente registrada a las 2:51:14 del 06 de mayo de 2021 con el Folio: 16926718

PRECIO: Mercado

Limitado \$

Vigencia: Diaria Semanal

Efectivo Estimado: \$ 1,030.85

Vigencia de la Orden: 06 may 2021

El precio es igual al precio de mercado.

Te recordamos que el precio por acción que aparece en pantalla es solo un precio de referencia que corresponde al precio del día anterior, por lo cual, no debe tomarse como definitivo. El precio definitivo se determinará de acuerdo al prospecto de información al público inversionista.

ENVIAR Cancelar

Ilustración 29 Alta de orden exitosa

Operación

ACCIONES

FONDOS DE INVERSIÓN

MOVIMIENTOS DEL DÍA

HISTÓRICO DE MOVIMIENTOS

Operación: Compra Venta

Emisora: BIMBO.A

Títulos:

Precio: Mercado Limitado \$ 0.000

Vigencia: Diaria Semanal

VALIDAR Limpiar

BIMBO.A

Precio Anterior \$ 40.93

Precio Último \$ 41.09

Diferencia \$ 0.16

Variación Porcentual 0.39%

Máx. / Mín. Hecho \$ 41.40 / \$ 40.80

ÓRDENES DEL DÍA

	Emisora	Títulos	Precio (\$)	Títulos Asignados	Títulos Pend.	Fecha y hora de Captura	Vigencia	
En proceso	Compra	BIMBO.A	25	MKT	0	25	06 may 2021 14:51:14	06 may 2021 15:05:00

Ilustración 30 Compra en proceso

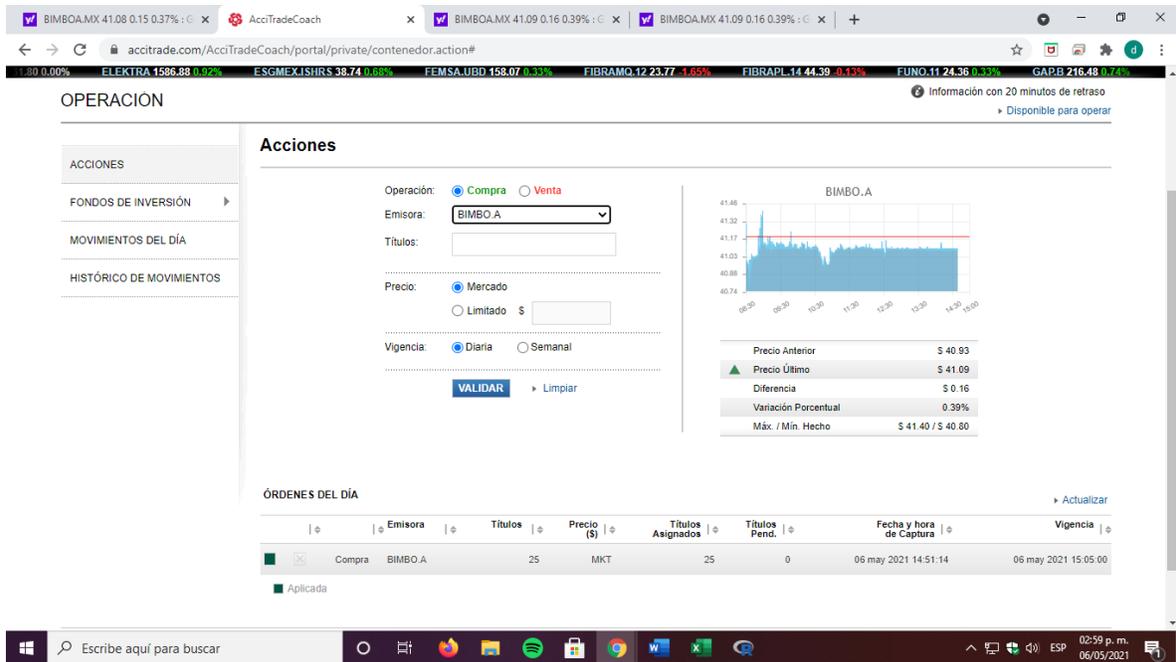


Ilustración 31 Compra aplicada

Al realizar esta compra muy cercana al cierre de la bolsa de valores, se pudieron comprar las acciones cuando bajaron un poco de precio, como se muestra en la Ilustración.



Ilustración 32 Precio de cierre al final del día financiero 6 de Mayo de 2021

## Valores de compra

Empresa	Acciones	Precio
BIMBO.A	25	41.00

Tabla 5 Valores de compra de 6 de Mayo de 2021

## 7 de Mayo de 2021

1. Se obtuvieron los datos acumulados de la semana actual (3 de Mayo de 2021) al 6 de Mayo de 2021 de las variables de entrada.

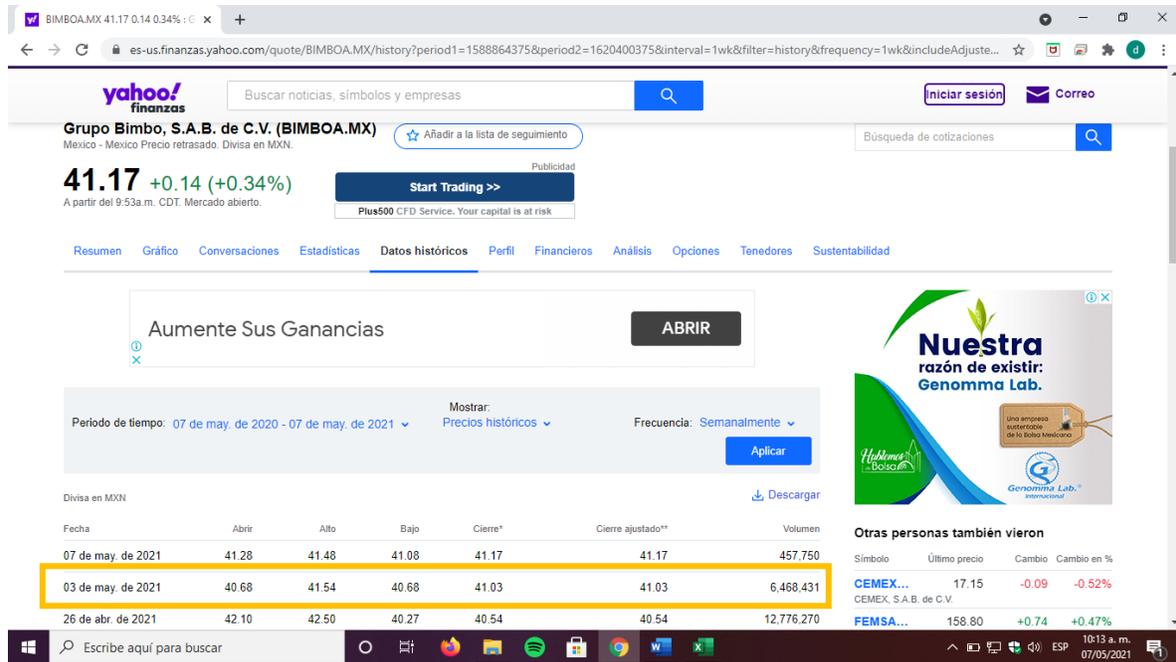


Ilustración 33 Datos de Grupo Bimbo semanales acumulados al 6 de Mayo de 2021

2. Estos datos se colocaron en un archivo .csv.

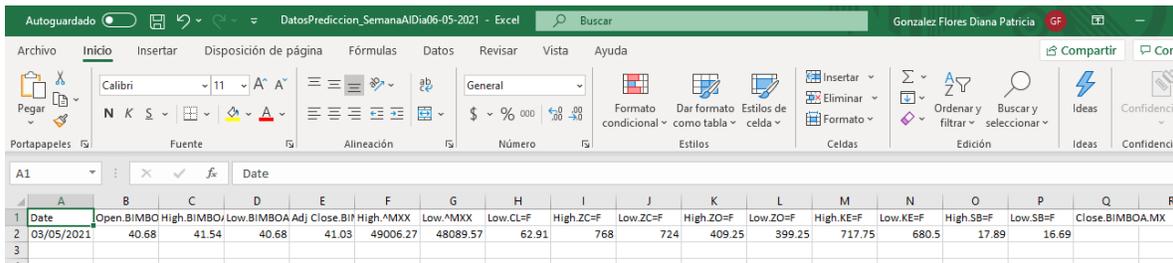


Ilustración 34 Datos de entrada para red neuronal

3. En R estos datos se ingresaron como entrada en la red neuronal y se pronosticó el precio de cierre correspondiente a estos datos. Esto se realiza de la misma manera en que se ejecutó la red neuronal con el conjunto de datos del día anterior.

```

> #predecir valores nuevos
> #lectura de datos
> datosPred <- read.csv("E:/8 Semestre/Mineria/Final/DatosPrediccion_SemanaAlDi$
> attach(datosPred)
The following objects are masked from datos:

  Adj.Close.BIMBOA.MX, Close.BIMBOA.MX, Date, High..MXX,
  High.BIMBOA.MX, High.KE.F, High.SB.F, High.ZC.F, High.ZO.F,
  Low..MXX, Low.BIMBOA.MX, Low.CL.F, Low.KE.F, Low.SB.F, Low.ZC.F,
  Low.ZO.F, Open.BIMBOA.MX

>
> datosPred[1,]
      Date Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX
1 03/05/2021          40.68           41.54           40.68           41.03
High..MXX Low..MXX Low.CL.F High.ZC.F Low.ZC.F High.ZO.F Low.ZO.F High.KE.F
1 49006.27 48089.57    62.91         768         724    409.25    399.25    717.75
Low.KE.F High.SB.F Low.SB.F Close.BIMBOA.MX
1    680.5    17.89    16.69                NA
>
> #quitar columna de fechas
> datosPred <- datosPred[,-1]
> datosPred[1,]
Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX High..MXX
1          40.68          41.54          40.68           41.03 49006.27
Low..MXX Low.CL.F High.ZC.F Low.ZC.F High.ZO.F Low.ZO.F High.KE.F Low.KE.F
1 48089.57    62.91         768         724    409.25    399.25    717.75    680.5
High.SB.F Low.SB.F Close.BIMBOA.MX
1    17.89    16.69                NA
>
> #normalización
> media_original <- attr(datos_nor,"scaled:center")
> desviacion_original <- attr(datos_nor,"scaled:scale")
> datos_pred_nor <- (datosPred-media_original)/desviacion_original
> datos_pred_nor
Open.BIMBOA.MX High.BIMBOA.MX Low.BIMBOA.MX Adj.Close.BIMBOA.MX High..MXX
1    0.6918856    0.5649524    0.9689963    0.8846297 1.835748
Low..MXX Low.CL.F High.ZC.F Low.ZC.F High.ZO.F Low.ZO.F High.KE.F Low.KE.F
1 1.884487 1.12115 4.478798 4.719473 2.812479 3.146098 2.965631 3.069269
High.SB.F Low.SB.F Close.BIMBOA.MX
1    2.38736 1.570323                NA
>
> valor_estimado_ModeloFinal <- (compute(red5, datos_pred_nor))$net.result
>
> #regresar a unidades originales
> valor_estimado_en_mxn_ModeloFinal <- (valor_estimado_ModeloFinal*desviacion)+$
> valor_estimado_en_mxn_ModeloFinal
      [,1]
[1,] 39.33545
> |

```

Ilustración 35 Pronóstico de precio de cierre en R

### Precio de cierre pronosticado: 39.33545 MXN

Con este valor se monitoreó el precio de las acciones de Grupo Bimbo durante el día financiero. Como se muestra en la imagen, el precio bajó un poco y se acercó un poco al precio pronosticado cerca de las 11:00 a.m. por lo que se decidió hacer una compra de más acciones.

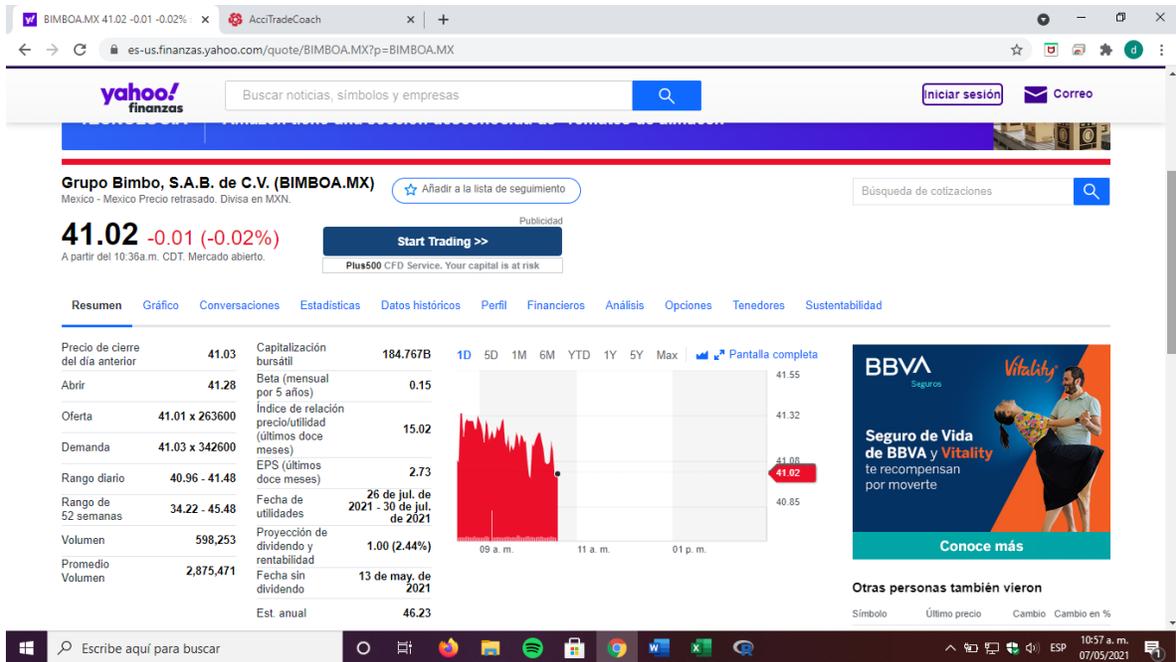


Ilustración 36 Monitoreo de precio de acciones

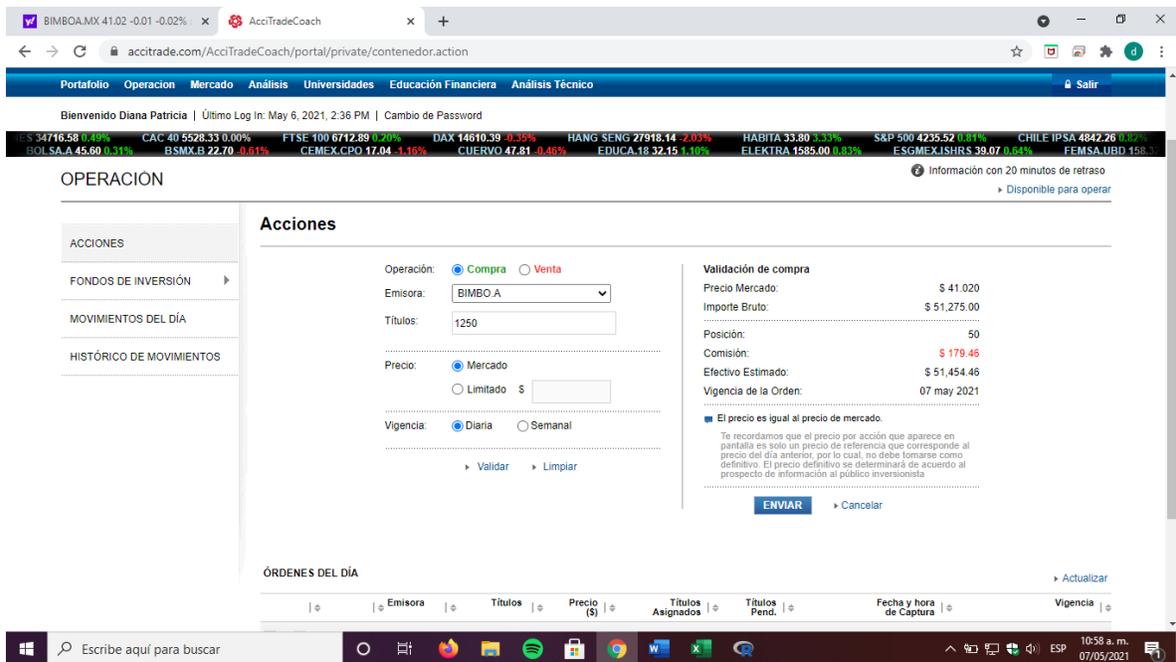


Ilustración 37 Compra de acciones 7 de Mayo de 2021

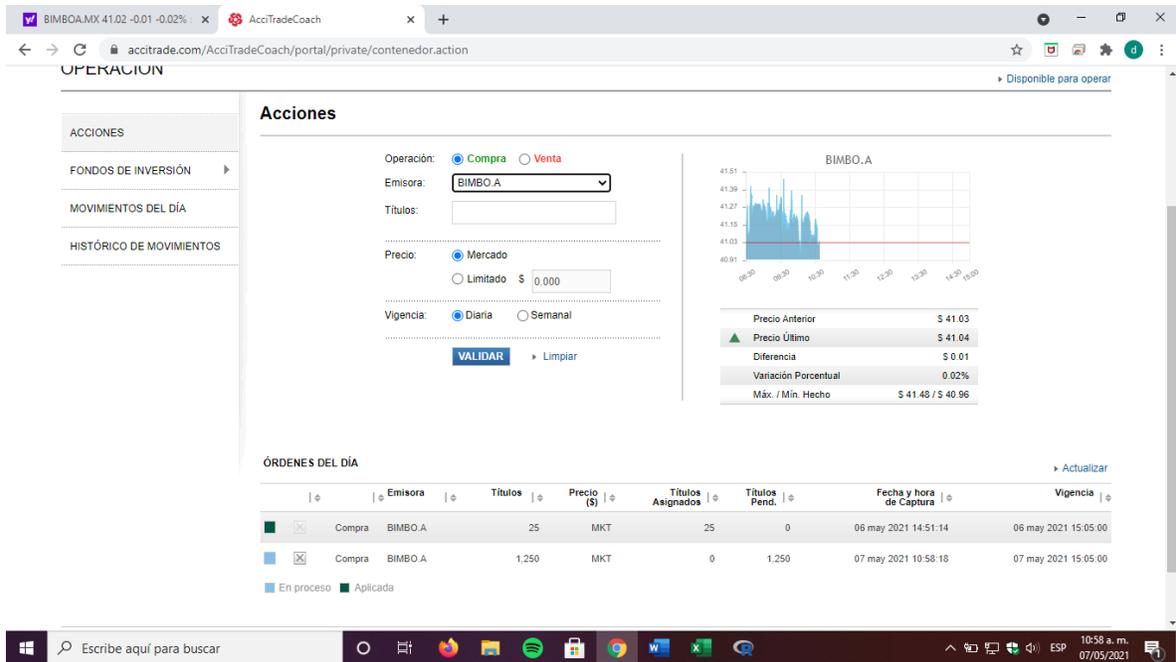


Ilustración 38 Compra en proceso

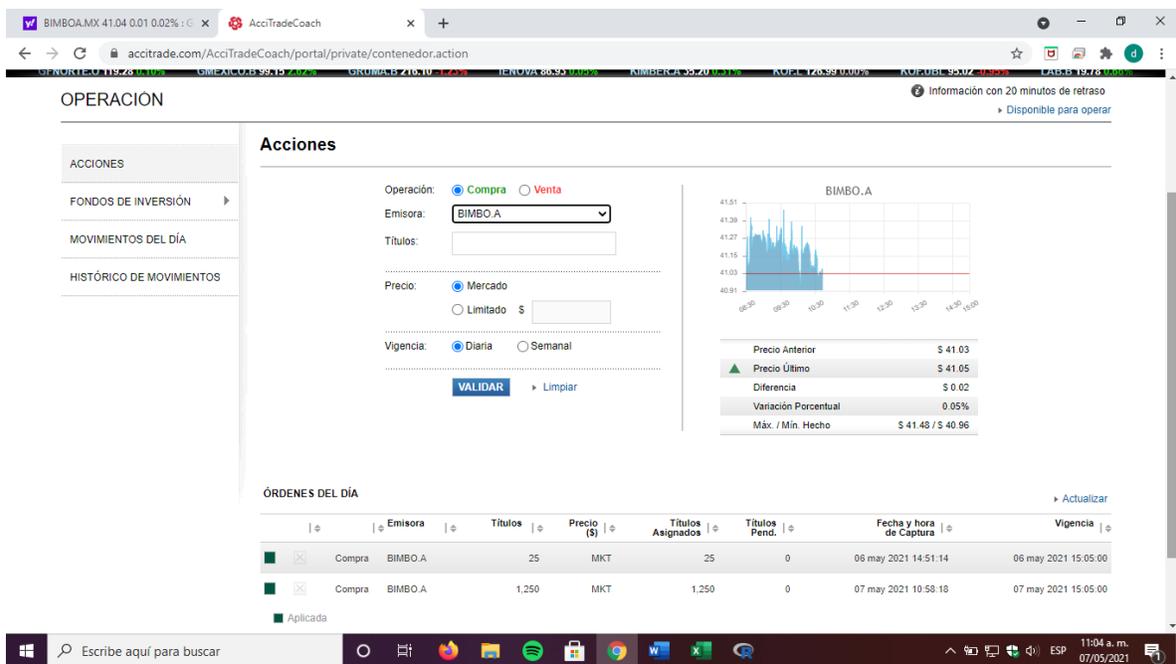


Ilustración 39 Compra aplicada

### Valores de compra

Empresa	Acciones	Precio
BIMBO.A	100	41.08
	100	41.08
	100	41.08

	100	41.09
	100	41.09
	100	41.09
	100	41.09
	100	41.09
	100	41.10
	100	41.10
	100	41.10
	100	41.10
	50	41.10
<b>Total</b>	<b>1250</b>	

Tabla 6 Valores de compra de 7 de Mayo de 2021

## Resultados

### Diferencias entre el precio de cierre pronosticado y el valor real

Después de terminar ambos días financieros, se registraron los precios de cierre reales para compararlos con los pronosticados, obteniendo lo siguiente:

	6 de Mayo de 2021	7 de Mayo de 2021
<b>Precio de cierre pronosticado</b>	39.35061 MXN	39.33545 MXN
<b>Precio de cierre real</b>	41.03 MXN	41.10 MXN
<b>ECM (Error cuadrado medio)</b>	2.82	3.11
<b>Eficiencia</b>	95.91%	95.71%

Tabla 7 Diferencias entre precio de cierre pronosticado y el precio real

ECM= (Precio de cierre real – Precio de cierre pronosticado) ^ 2

$$\text{Eficiencia} = 100 - \frac{(|\text{Precio de cierre real} - \text{Precio de cierre pronosticado}|) * 100}{\text{Precio de cierre real}}$$

### Ganancias y Pérdidas

Después de dos semanas de haber realizado las compras de acciones, el 21 de Mayo de 2021 se monitoreó el precio de las mismas durante el día financiero, el precio fue el siguiente:

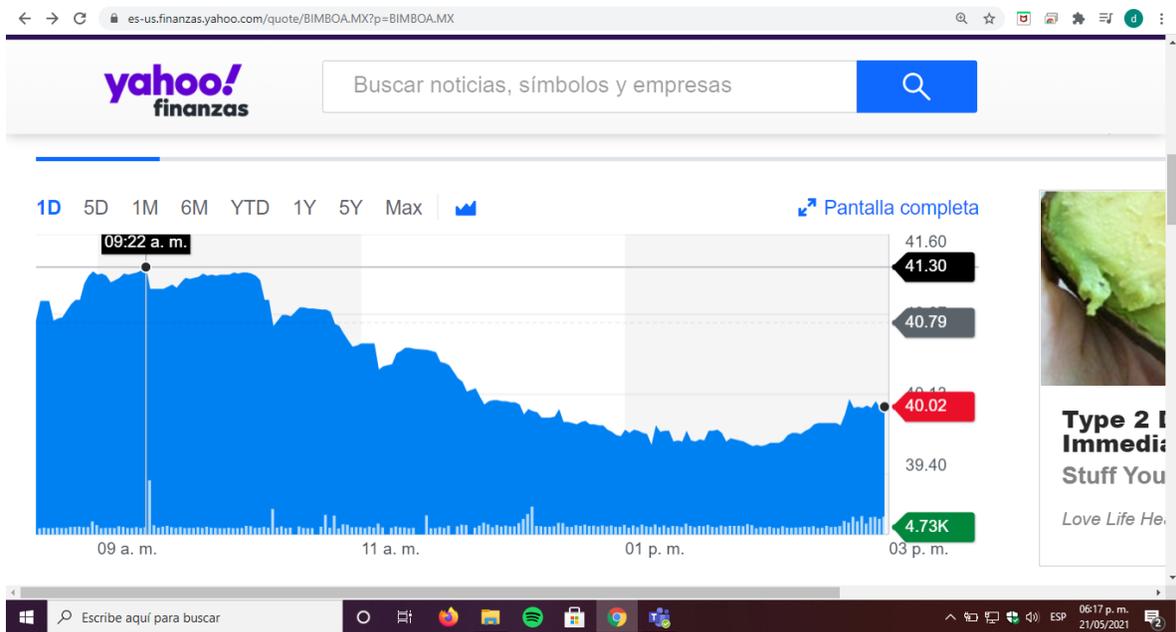


Ilustración 40 Precio de las acciones de Grupo Bimbo el 21 de Mayo de 2021

Como se puede observar en el gráfico, a las 9:22 am su precio llegó a 41.30 MXN y su precio de cierre fue 40.02 MXN.

En AcciTrade Coach, con el precio de cierre, se obtuvieron los siguientes valores:

The screenshot shows the AcciTrade Coach interface. The 'PORTAFOLIO' section is active, displaying a 'Detalle de Posición' for BIMBOA.MX. The total portfolio value is \$2,008,059.12. The table below provides a breakdown of the investment in BIMBOA.MX and other stocks.

RENDA VARIABLE									
Emisora	Cantidad	Costo (\$)	Inversión (\$)	Precio (\$)	Imp. Val. (\$)	Cambio (%)	Cambio (\$)	Fec. Ult. Mov.	Total (%)
BIMBOA	1,300	41.12	53,456.50	40.02	52,026.00	-2.68	-1,430.50	07-may-2021 10:59	2.59
CUERVO	22	46.80	1,029.60	47.89	1,053.58	2.33	23.98	06-abr-2021 11:58	0.05
LIVERPOL.C1	14	74.29	1,040.06	87.67	1,227.38	18.01	187.32	06-abr-2021 12:00	0.06
WALMEX	16	65.59	1,049.44	62.88	1,006.08	-4.13	-43.36	06-abr-2021 11:57	0.05
<b>TOTAL</b>			56,575.60		55,313.04				2.75

SOCIEDADES DE INVERSIÓN									
Emisora	Cantidad	Costo (\$)	Inversión (\$)	Precio (\$)	Imp. Val. (\$)	Cambio (%)	Cambio (\$)	Fec. Ult. Mov.	Total (%)
<b>TOTAL</b>			0.00		0.00				0.00

Ilustración 41 Valores de las acciones en AcciTrade Coach

Los resultados al 21 de Mayo del 2021 son los siguientes:

- Cantidad de acciones: 1300
- Costo de compra promedio: 41.12 MXN

- Inversión: 53,456.50 MXN

Hora	Precio	Valor de las acciones	Cambio
9:22 a.m.	41.30 MXN	53,690.00	233.50
2:59 p.m.	40.02 MXN	52,026.00	-1430.50

Tabla 8 Ganancias y Pérdidas

Inversión = Cantidad de acciones \* Costo de compra promedio

Valor de las acciones = Precio \* Cantidad de acciones

Cambio = Valor de las acciones - Inversión

## Conclusiones

Al pronosticar el precio de cierre de las acciones de Grupo Bimbo, se tuvo un porcentaje de eficiencia cercano al 96%, lo cual es bastante alto, pero como se pronosticaron precios semanales, no se puede asegurar que la semana posterior a la que se pronosticó, los precios cambien a nuestro favor y se obtengan ganancias.

Como se puede observar en los resultados, el precio de las acciones de Grupo Bimbo varía incluso a lo largo del día financiero, pues a las 9:22 a.m. se obtuvo una pequeña ganancia, pero al cierre, se tenía una pérdida.

Entonces, se puede concluir que para tomar decisiones de compra y venta de acciones, se requiere más conocimiento que solo el pronóstico del precio de cierre de las acciones.

También se debe tomar en cuenta que otras variables pueden influir en el precio de las acciones de Grupo Bimbo, como cambios organizacionales, ventas semanales, número de empleados, número de fábricas etc. Además, respecto al precio de las materias primas, se utilizaron los precios en la bolsa de valores para dar una perspectiva general del mercado, pero Grupo Bimbo tiene sus proveedores y puede que compren estas materias primas a un precio más accesible por la cantidad de materias que necesitan para sus operaciones. En este caso no se tiene acceso a estas variables, ya que es información privada de Grupo Bimbo o solo está disponible para sus inversionistas.

Tomando en cuenta lo anterior, se considera importante que en este tipo de proyectos se cuente con un equipo multidisciplinario, ya que es importante tener un buen conocimiento del tema en el que se van a aplicar técnicas de minería de datos, en este caso sería apropiado contar con la opinión de un analista financiero y/o algún experto en temas de economía, ya que podrían dar una mejor perspectiva sobre las acciones y que variables específicamente pueden afectar el precio de las acciones y donde se pueden encontrar estos datos.

Adicionalmente, un experto en temas financieros podría tomar mejores decisiones de compra y venta de acciones, ya que tienen más experiencia y le pueden dar más significado al valor pronosticado, ya que en este caso solo se compraron acciones cuando el precio se acercaba un poco al precio pronosticado y como se puede observar en los resultados, no se obtuvieron muchas ganancias. El valor pronosticado solo es una referencia, ya que puede que existan más factores que sean necesarios de tomar en cuenta al realizar compra/venta de acciones en la bolsa de valores.

Al día de finalización del proyecto, solo pasaron dos semanas desde la compra de acciones, tal vez dentro de un mes o más tiempo pueda cambiar el precio de las acciones y se obtenga una ganancia o pérdida mayor. Se podrían ir pronosticando los valores del precio de las acciones del Grupo Bimbo durante varios meses para ir agregando estos nuevos valores al conjunto de entrenamiento para ir mejorando la red neuronal, de esta manera se podría mejorar su efectividad y con un experto en temas de la bolsa de valores, se podría llegar a utilizar en el mundo real.

El presente proyecto solo consistió en la aplicación de técnicas de minería de datos en compra de acciones de la bolsa de valores, pero el campo de aplicación de la minería de datos es inmenso, así que se puede utilizar en muchos campos de estudio para solucionar diversos problemas y para proponer nuevas soluciones que pueden beneficiar a diversos sectores de la población.

## Propuestas

Con la minería de datos se pueden lograr una infinidad de cosas, ya que ésta se puede aplicar en diversas áreas como:

- Marketing
- Comercio
- Finanzas
- Salud
- Medios de comunicación
- Redes
- Ingeniería
- Educación
- Y muchas áreas más...

Específicamente este tipo de solución (red neuronal con algoritmo backpropagation) se puede aplicar en diversos problemas de la vida cotidiana como:

### Entidades financieras

Las entidades financieras otorgan créditos según ciertas características como ingresos, egresos, garantías, etc. De esta manera se crea una evaluación del usuario que pide el crédito para establecer si se autoriza el crédito y el monto de éste. Esta evaluación y asignación de crédito, es realizada por analistas financieros,

haciendo uso de su propia experiencia. De esta manera no se puede asegurar que estas evaluaciones sean correctas ya que no se puede saber el nivel de morosidad del solicitante (deudor que no paga al vencimiento de su obligación) y esto puede afectar a la entidad financiera generando pérdidas de capital. Con minería de datos se podría desarrollar una herramienta para hacer estas evaluaciones y clasificar a los solicitantes según su nivel de morosidad. De esta manera se podrían hacer decisiones informadas con respecto a la autorización de créditos.

Cornejo y Quispe (2011) describen una aplicación del algoritmo Backpropagation para calcular el nivel de morosidad de solicitantes para poder analizar la asignación del crédito.

### **Contaminación**

Con el paso del tiempo, el planeta se ve cada vez más afectado por la contaminación, esto nos afecta a todos por lo que es importante plantear estrategias para mitigar este tipo de afectaciones.

Por ejemplo, se podría utilizar un conjunto de datos relacionados con factores ambientales para predecir los índices de contaminación atmosférica para tomar acciones y contrarrestar sus efectos negativos en nuestro entorno.

Díaz (2014) muestra el desarrollo de un sistema integral bajo el enfoque de minería de datos y redes neuronales para la predicción y control de la contaminación atmosférica por pm10 en la ciudad de Chimbote, Perú.

### **Industria**

Este tipo de redes neuronales se pueden utilizar en diversas áreas de la industria, desde cadenas de suministro, procesos, gestión, calidad, etc. Con estas aplicaciones se podrían mejorar diversos aspectos involucrados en las empresas. De esta manera se podrían incrementar las ganancias, disminuir riesgos, disminuir pérdidas, entre muchas cosas más.

Por ejemplo, González (2006) describe el uso de técnicas de minería de datos para el modelado del proceso industrial de una línea de acero galvanizado. De esta manera se predicen las propiedades mecánicas de las bobinas de acero galvanizado y la velocidad de la banda dentro del horno de proceso para hacer mejoras en los sistemas de control.

### **Salud**

En el ámbito de salud se pueden proponer una infinidad de soluciones que impliquen el uso de minería de datos. Desde determinar causas de ciertas enfermedades o a identificar poblaciones de riesgo hasta la gestión de sistemas de salud.

Suca, Córdoba, Condori, Cayra y Sulla (2016) compararon algoritmos de clasificación para la predicción de casos de obesidad infantil. Aquí se utilizaron

algunas técnicas de minería de datos para que modelos puedan aprender y predecir de acuerdo a un conjunto de datos si un niño tendrá obesidad. Con esta solución se ayudaría a las entidades que evalúan a nivel mundial los índices de obesidad infantil a tomar acción para prevenir este tipo de situaciones y mejorar la salud de la niñez.

## Código en R

```
1. #abrir librerías necesarias
2. library (grid)
3. library (MASS)
4. library (neuralnet)
5. library(e1071)
6.
7. #para limpiar memoria
8. rm(list=ls())
9.
10. #inicializar generador de numeros aleatorios
11. set.seed(42)
12.
13. #lectura de datos
14. datos <- read.csv("E:/8 Semestre/Mineria/Parcial
15. 3/Bimbo/conjuntoDeDatos_BIMBOA.MX.csv",header=TRUE)
16. attach(datos)
17.
18. datos[1,]
19.
20. #quitar columna de fechas
21. datos <- datos[,-1]
22. datos[1,]
23.
24. #correlación
25. datos_cor <- cor(datos)
26.
27. #quitar columnas con baja correlación
28. datos2 <- datos[,-5]
29. datos2 <- datos2[,-7]
30. datos2 <- datos2[,-7]
31. datos2 <- datos2[,-7]
32. datos2[1,]
33.
34. #normalización
35. datos_nor <- scale(datos2)
36. datos_nor
37.
38. #separar conjunto de entrenamiento y prueba
39. tamaño_muestra <- (trunc(0.7*nrow(datos_nor)))
40. muestra <- sample(1:nrow(datos_nor),tamaño_muestra)
41. entrena <- datos_nor[muestra,]
42. prueba <- datos_nor[-muestra,]
43. entrena
44. prueba
45.
46. # ----- MODELO 1 -----
47. #red de 5 neuronas
48. red5 <-
49. neuralnet(Close.BIMBOA.MX~Open.BIMBOA.MX+High.BIMBOA.MX+Low.BIMBOA.MX+Adj.Close.BIMBOA
50. .MX+High.MX+Low.MX+Low.CL.F+High.ZC.F+Low.ZC.F+High.ZO.F+Low.ZO.F+High.KE.F+Low.KE
51. .F+High.SB.F+Low.SB.F, entrena, hidden=5, threshold=0.01)
52. plot(red5)
```

```

51. valor_estimado_Modelo1 <- (compute(red5, prueba))$net.result
52.
53. #regresar a unidades originales
54. media <- mean (datos2[,16])
55. desviacion <- sd (datos2[,16])
56.
57. valor_estimado_en_mxn_Modelo1 <- (valor_estimado_Modelo1*desviacion)+media
58.
59. #error
60.
61. diferencias <- datos2[-muestra,16]-valor_estimado_en_mxn_Modelo1
62.
63. dif_cuad <- diferencias^2
64.
65. ECM_Modelo1 <- mean(dif_cuad)
66. ECM_Modelo1
67.
68. # ----- MODELO 2 -----
69.
70. #red de 10 neuronas
71. red10 <-
  neuralnet(Close.BIMBOA.MX~Open.BIMBOA.MX+High.BIMBOA.MX+Low.BIMBOA.MX+Adj.Close.BIMBOA
  .MX+High..MX+Low..MX+Low.CL.F+High.ZC.F+Low.ZC.F+High.ZO.F+Low.ZO.F+High.KE.F+Low.KE
  .F+High.SB.F+Low.SB.F, entrena, hidden=10, threshold=0.01)
72. plot(red10)
73.
74. valor_estimado_Modelo2 <- (compute(red10, prueba))$net.result
75.
76. #regresar a unidades originales
77. valor_estimado_en_mxn_Modelo2 <- (valor_estimado_Modelo2*desviacion)+media
78.
79. #error
80.
81. diferencias <- datos2[-muestra,16]-valor_estimado_en_mxn_Modelo2
82.
83. dif_cuad <- diferencias^2
84.
85. ECM_Modelo2 <- mean(dif_cuad)
86.
87. ECM_Modelo2
88.
89. # ----- MODELO 3 -----
90.
91. #red de 15 neuronas
92. red15 <-
  neuralnet(Close.BIMBOA.MX~Open.BIMBOA.MX+High.BIMBOA.MX+Low.BIMBOA.MX+Adj.Close.BIMBOA
  .MX+High..MX+Low..MX+Low.CL.F+High.ZC.F+Low.ZC.F+High.ZO.F+Low.ZO.F+High.KE.F+Low.KE
  .F+High.SB.F+Low.SB.F, entrena, hidden=15, threshold=0.01)
93. plot(red15)
94.
95. valor_estimado_Modelo3 <- (compute(red15, prueba))$net.result
96.
97. #regresar a unidades originales
98. valor_estimado_en_mxn_Modelo3 <- (valor_estimado_Modelo3*desviacion)+media
99.
100. #error
101.
102. diferencias <- datos2[-muestra,16]-valor_estimado_en_mxn_Modelo3
103.
104. dif_cuad <- diferencias^2
105.
106. ECM_Modelo3 <- mean(dif_cuad)
107.
108. ECM_Modelo3
109.

```

```

110. #errores
111. ECM_Modelo1
112. ECM_Modelo2
113. ECM_Modelo3
114.
115. #en porcentajes
116. #usar modelo con menor error
117. diferencias <- datos2[-muestra,16]-valor_estimado_en_mxn_Modelo1
118.
119. porcentajes <-abs(diferencias)*100/datos2[-muestra,16]
120.
121. promedio_porcentajes <- mean(porcentajes)
122.
123. eficiencia <- 100-promedio_porcentajes
124.
125. eficiencia
126.
127. #predecir valores nuevos
128. #lectura de datos
129. datosPred <- read.csv("E:/8 Semestre/Mineria/Final/DatosPrediccion_SemanaAlDia06-05-
2021.csv",header=TRUE)
130. attach(datosPred)
131.
132. datosPred[1,]
133.
134. #quitar columna de fechas
135. datosPred <-datosPred[,-1]
136. datosPred[1,]
137.
138. #normalización
139. media_original <- attr(datos_nor,"scaled:center")
140. desviacion_original <- attr(datos_nor,"scaled:scale")
141. datos_pred_nor <- (datosPred-media_original)/desviacion_original
142. datos_pred_nor
143.
144. valor_estimado_ModeloFinal <- (compute(red5, datos_pred_nor))$net.result
145.
146. #regresar a unidades originales
147. valor_estimado_en_mxn_ModeloFinal <- (valor_estimado_ModeloFinal*desviacion)+media
148. valor_estimado_en_mxn_ModeloFinal

```

## Bibliografía

- Accitrade Coach. (s.f.). *Preguntas Frecuentes*. Obtenido de Accitrade Coach: <https://www.accitrade.com/AcciTradeCoach/portal/public/ayuda.do>
- Beltrán Marínez, B. (2018). *Minería de datos*. Puebla: Facultad de Ciencias de la Computacion, Benemérita Universidad Autónoma de Puebla.
- Calvo, D. (8 de Diciembre de 2018). *Perceptrón – Red neuronal*. Obtenido de Diego Calvo: <https://www.diegocalvo.es/perceptron/>
- Caurin, J. (2019). *¿Cómo afecta el IPC a la economía de las personas y empresas?* Obtenido de Economía Simple: <https://www.economiasimple.net/como-afecta-el-ipc-a-la-economia-de-las-personas-y-empresas.html>

- Cornejo Ruiz, D., & Quispe Gavino, G. (2011). Aplicación del algoritmo Backpropagation de redes neuronales para determinar los niveles de morosidad en los alumnos de la Universidad Peruana Unión. *Revista de Investigación Business Intelligence*, 21-31.
- Díaz Tello, S. (2014). *Sistema integral bajo el enfoque de minería de datos y redes neuronales para la predicción y control de la contaminación atmosférica por PM10 en la ciudad de Chimbote*. Chimbote: Universidad Nacional del Santa.
- González Marcos, A. (2006). *Desarrollo de técnicas de minería de datos en procesos industriales: Modelización en líneas de producción de acero*. Logroño: Universidad de la Rioja.
- Grupo Bimbo. (2021). *Inicio*. Obtenido de Grupo Bimbo: <https://www.grupobimbo.com/es>
- Hernández Orallo, J., Ramírez Quintana, M. J., & Ferri Ramírez, C. (2004). *Introducción a la minería de datos*. Madrid: Pearson Educación.
- Jiménez Bermejo, D. (2015). *Acción*. Obtenido de Economipedia: <https://economipedia.com/definiciones/accion.html>
- Molina Félix, L. C. (2002). *Data mining: torturando a los datos hasta que confiesen*. Barcelona: Departamento de Lenguajes y Sistemas Informáticos, Universidad Politécnica de Cataluña.
- Roldán, P. N. (2020). *Bolsa de valores*. Obtenido de Economipedia: <https://economipedia.com/definiciones/bolsa-de-valores.html>
- Suca, C., Córdoba, A., Condori, A., Cayra, J., & Sulla, J. (2016). *Comparación de algoritmos de clasificación para la predicción de casos de obesidad infantil*. Arequipa: Universidad Nacional de San Agustín.
- Toral Barrera, J. A. (2009). *Redes Neuronales*. Obtenido de Universidad de Guadalajara: [http://www.cucei.udg.mx/sites/default/files/pdf/toral\\_barrera\\_jamie\\_areli.pdf](http://www.cucei.udg.mx/sites/default/files/pdf/toral_barrera_jamie_areli.pdf)
- Valcárcel Asencios, V. (2004). Data Mining y el descubrimiento del conocimiento. *Industrial Data vol. 7, núm. 2*, 83-86.
- Vázquez Burguillo, R. (2015). *IPC mexicano*. Obtenido de Economipedia: <https://economipedia.com/definiciones/ipc-mexicano.html>