Predictor de las series de tiempo del dólar en COP (pesos colombianos)

Juan David Peralta Rodriguez

Universidad Nacional de Colombia

La empresa importamos S.A se dedica al negocio de la importación de productos, específicamente se dedica a la importación de mercancía desde estados unidos hacia Colombia. Estos productos son mayormente vendidos por internet y para pequeños distribuidores en Colombia, es importante entender que este es un negocio de oportunidades, por lo tanto están a la expectativa de como el dólar puede fluctuar y como esto puede significar una importante pérdida o ganancia.

Se espera que pueda comprar los productos cuando es más barata la tasa de cambio. En el siguiente grafico vemos como una fluctuación de 2 % puede ser muy normal en un rango de tres meses por lo tanto esta puede significar una pérdida de capital, esta empresa trabaja en promedio con un capital mensual de 50 millones de pesos. Por lo tanto puede significar una pérdida de más de 1 millón de pesos, y un mal día puede significar peores pérdidas. Lo podemos ver en el siguiente grafico



Esta es una ventana de 6 meses de cómo ha fluctuado el dólar. Y en café está el dato de la volatilidad de los últimos 20 días el cual nos habla de lo peligroso que puede ser comprar en un mal momento.

El negocio de las importaciones puede estar demasiado atraído a estar siempre un paso adelante conociendo como fluctúa la tasa de cambio, sabemos que esto es casi imposible demasiados factores macroeconómicos están involucrados. Pero aquí es cuando el maching learning puede ayudarnos un poco pues conocemos los históricos del dólar y este sistema no necesita una teoría clara de cómo funciona el mercado mundial, simplemente necesitamos la información y una computadora.

Vamos a mostrar la tasa de cambio de otras dos economías Latinoamericanas, para mostrar cómo podemos ver dos casos totalmente contrarios y como estas razones de cambio pueden ser totalmente volátiles.

Mostramos el valor del dólar en soles peruanos, la moneda nacional de Perú, desde 1997 hasta la fecha:

Este grafico se ve muy estable, bastante parecido a un sistema estacionario.

Ahora mostraremos por el contrario una economía que su tasa de cambio ha aumentado a una velocidad alarmante, este es la tasa de cambio para Venezuela desde el año 2010 hasta marzo del 2021:

Vemos como aumenta a una velocidad demasiado grande en un rango de tiempo muy corto, esto sería catastrófico para un negocio como el de IMPORTADORA S.A. Por lo tanto vemos como tener alguna certeza o al menos un pronóstico de cómo se va a comportar el dólar será muy importante.

IMPORTAMOS S.A es una empresa dedicada a la importación de mercancía desde estados unidos a Colombia cuenta con la capacidad de traer containers con mercancía, esta mercancía la compra al por mayor en EE.UU tratando de bajar los costos tanto en venta como en trasporte.

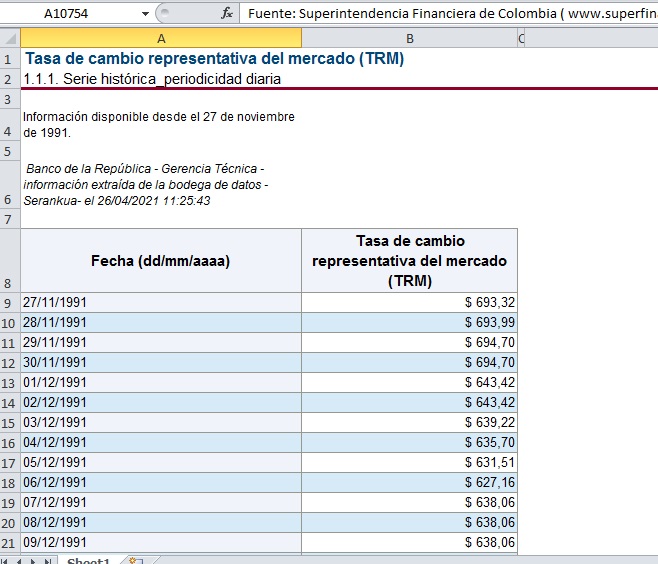
Con la mercancía comprada venden los productos algunos a distribuidores en Colombia así como ellos poseen páginas de internet propias en las cuales ofrecen los productos. Las ganancias de la empresa están dictaminadas por tres factores muy importantes:

1. La venta de los productos dependiendo del mercado será más beneficiosa o no, es muy importante entender que la empresa nunca debe perder competitividad por lo tanto hacer los productos más costosos es una opción que debe tomarse como último recurso.
2. Cantidad de producto dependiendo de la situación esto se maneja por medio de decisiones gerenciales de la empresa.
3. Por último estará la tasa de cambio US-COP, este es el aspecto en el que más estamos interesados en este proyecto. Simplemente tener un dólar 2% más barato o más caro puede implicar perdidas millonarias.

Data:

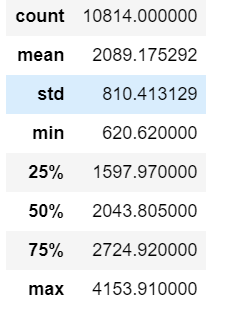
La data de nuestro modelo es una tabla con el día mes y año, en otra columna tenemos a el precio de 1 dólar en pesos colombianos. La fuente de la información es la Superintendencia Financiera de Colombia ( [www.superfinanciera.gov.co](http://www.superfinanciera.gov.co)).

A continuación mostramos como viene la información y este es el link donde se puede conseguir <https://www.banrep.gov.co/es/estadisticas/trm>



La información es conseguida del banco de la república de Colombia esta data se actualiza de forma diaria y es el valor más confiable que se puede conseguir. La encontramos de forma completa y todos datos utilizables.

Como la información que conseguimos es confiable y correcta, no necesitamos hacer un proceso de selección o limpieza, por lo tanto simplemente se hace las estadísticas de la data, encontramos los siguientes valores característicos:



Encontramos una base de 10814 valores estos son los datos desde el 27 de noviembre de 1991 hasta la fecha actual 1 datos por día. También encontramos un mínimo de $620.62 el 21 de diciembre de 1991 y un máximo de $4153.91 el 20 de febrero de 2020. También tenemos los cuartiles :

* 1° cuartil $1597
* 2° cuartil $2043
* 3° cuartil $2724

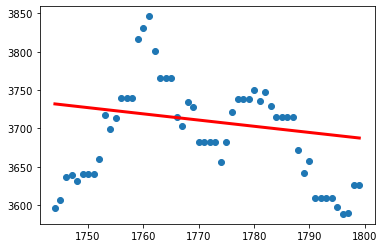
Para la selección de la data nosotros utilizamos el 80 % para entrenar el modelo y 20 % para poder testearlo como sabemos en el maching learning tenemos que darle una cantidad de información para que este pueda entrenarse y otra cantidad de información para que este evalué la efectividad del modelo en este caso se utiliza los primeros datos que teníamos en este modelo genere un algoritmo y datos más recientes para que podamos evaluar al modelo. Como tenemos una serie de tiempo no podemos alterar la el orden de los datos.

Para el modelo probamos varias alternativas una de ellas fue la red neuronal bastante atractiva por su precisión y simplicidad pero se tuvo problemas con este modelo, la cantidad de neuronas necesarias era demasiado grande igual que el tamaño de nuestros datos y si dado el caso quisiéramos actualizar los datos sería necesaria la creación de neuronas. Además los pesos son principales en el proyecto es necesario que estén separados cada uno de los datos en una neurona y darle un peso más alto conforme avanzamos en la fecha. Probamos otros modelos igual de atractivos que las redes neuronales pero por su complejidad fue necesario entrar a evaluar cuál era el objetivo del proyecto.

Como nada ayudaba en la precisión y significaba un gasto muy grande. También un dato importante es que los datos no pueden ser divididos aleatoriamente pues pierde la lógica que existe en estos datos, la ubicación temporal.

Optamos por una simple regresión polinomial encontramos que esta capacidad de predicción no era demasiado grande pues no podemos predecir fechas alejadas del punto de cierre de los datos esta si tiene un buena precisión en datos seguidos al punto de corte.

Fue necesario también hacer un arreglo de pesos para lograr que el modelo entendiera que entre más cerca estamos de la fecha de corte es más importante este dato. Para esto a medida que vamos avanzando en el dato multiplicamos su ubicación y dividimos entre la cantidad de datos que tenemos. El cual hace que tenga un factor de peso muy grande cerca de la fecha de cierre de los datos, tener en cuenta que estos datos se actualizan diariamente.



Encontramos esta regresión polinomial que cuadra muy bien con la dinámica que tiene el dólar, tiene unos valores muy cercanos a los valores que fueron dados por el banco de la república. Y principal mente predice que estamos en una situación que estaba bajando nuevamente el dólar.

Tenemos índice de error promedio para estos datos de 1.43 %

Conclusiones

Encontramos que la regresión polinomial es una buena alternativa para lograr encontrar una aproximación del valor del dólar, siendo preciso para fechas cercanas a la última fecha de corte en los datos de prueba del algoritmo.

Podemos estar cercanos a ver como es el comportamiento de un rango de días, no muchos.

Decidir hacer una compra es ideal teniendo un comportamiento descendente en el modelo pero es perfecto cuando comenzamos a ver nuevamente un comportamiento ascendente previniendo un aumento del dólar en los próximos días.

Vemos que tenemos muchísimas alternativas para hacer este algoritmo más preciso, una de las más interesante es utilizar más parámetros, claramente estos ligados a variables macroeconómicas, un ejemplo es el valor del barril de petróleo el cual va muy ligado al mercado mundial y al valor de las divisas.

Aunque la moneda fluctúa siempre el dólar aumenta el precio por lo tanto cualquier modelo o la mayoría tendrá a predecir un aumento suave en el precio, conforme pasa el tiempo.

Apéndice

**import** pandas **as** pd

**import** numpy **as** np

df **=** pd.read\_excel(io **=** "\\Users\CPE\Documents\ciencia de datos\\USD-COP.xlsx", sheet\_name**=**"Hoja1")

print(df)

df.shape

**from** sklearn.model\_selection **import** train\_test\_split

**from** sklearn.linear\_model **import** LinearRegression

y\_fx**=**df['Fecha (dd/mm/aaaa)'].reset\_index(drop**=True**)

y\_fx2**=**np.array(y\_fx).reshape(**-**1,1)

y\_fx.reset\_index(drop**=True**)

y\_fx

y\_fx**=**df['Tasa de cambio representativa del mercado (TRM)'].reset\_index(drop**=True**)

y\_fx

y\_train**=**y\_fx[0:10740]

y\_train

x\_train**=**y\_train.index

x\_train**=**np.array(x\_train).reshape(**-**1,1)

x\_train

y\_train**=**np.array(y\_train).reshape(**-**1,1)

y\_train

model **=** LinearRegression()

model.fit(x\_train, y\_train)

y\_test**=**np.array(y\_test).reshape(**-**1,1)

y\_test

x\_real**=**np.arange(1744,1800)

x\_real

x\_real**=**np.array(x\_real).reshape(**-**1,1)

y\_predict **=** model.predict(x\_real)

print(y\_predict)

y\_todo**=**y\_fx

x\_todo**=**y\_fx.index

x\_todo **=** np.array(x\_todo).reshape(**-**1,1)

model **=** LinearRegression()

model.fit(x\_todo, y\_todo)

X\_test\_poli **=** poli\_reg.fit\_transform(x\_real)

pr **=** linear\_model.LinearRegression()

pr.fit(X\_train\_poli, y\_todo)

Y\_pred\_pr **=** pr.predict(X\_test\_poli)

​

y\_todo**=**np.array(y\_todo).reshape(**-**1,1)

y\_todo

peso**=**[0]**\***len(y\_todo)

**for** i **in** range (0,len(y\_todo)):

peso[i]**=**i**\*\***5**/**len(y\_todo)

​

**from** sklearn.preprocessing **import** PolynomialFeatures

**from** sklearn **import** datasets, linear\_model

poli\_reg **=** PolynomialFeatures(degree **=** 2)

X\_train\_poli **=** poli\_reg.fit\_transform(x\_todo)

X\_test\_poli **=** poli\_reg.fit\_transform(x\_real)

pr **=** linear\_model.LinearRegression()

pr.fit(X\_train\_poli, y\_todo,peso)

Y\_pred\_pr **=** pr.predict(X\_test\_poli)

print(Y\_pred\_pr)

y\_test\_p **=** np.array([3595.57, 3606.42,3636.26,3639.12, 3630.81,3640.07,3640.07,3640.07,3659.62, 3717.46,3699.74,3712.89,3740.14,3740.14,3740.14,3816.65,3831.35,3846.28,3800.33,3765.33,3765.33,3765.33,3714.94,3703.20,3734.09,3728.09,3682.84,3682.84,3682.84,3682.84,3655.74,3682.66,3721.57,3738.19,3738.19,3738.19,3750.66,3735.41,3747.48,3729.02,3715.28,3715.28,3715.28,3715.28,3671.38,3642.29,3657.41,3609.20,3609.20,3609.20,3609.20,3597.18,3588.41,3589.86,3626.02,3626.02])

**import** matplotlib.pyplot **as** plot

plt.scatter(x\_real, y\_test\_p)

plt.plot(x\_real,Y\_pred\_pr , color**=**'red', linewidth**=**3)

plt.show()

Referencias:

# Machine Learning-based USD/PKR Exchange Rate Forecasting Using Sentiment Analysis of Twitter Data. DOI:10.32604/cmc.2021.015872

# <https://towardsdatascience.com/exchange-rate-predictiontime-series-forecasting-with-arima-27260faafcdc>

Regresión Polinómica en Python con scikit-learn. <https://www.iartificial.net/regresion-polinomica-en-python-con-scikit-learn/>

First neuran network for beginners explained(with code). Arnx , Arthur https://towardsdatascience.com/first-neural-network-for-beginners-explained-with-code-4cfd37e06eaf