

MODELO DE RECOMENDACIONES SOBRE EL PRECIO FUTURO DE UNA DIVISA

Proyecto Integrador
Maestría en Ciencias de los Datos y Analítica
Universidad EAFIT – Mayo de 2020

Álvaro Villa Vélez

Luis Vesga Vesga

Jorge Luis Rentería Roa

Edgar Leandro Jiménez Jaimes

Santiago Echeverri Calderón

Contexto

En la última década la acumulación de datos financieros ha crecido de una forma sin precedentes. El rápido crecimiento del volumen de datos representa un reto para analizarlos.

Objetivo

Desarrollar y evaluar un modelo de **clasificación**, que genere recomendaciones de **compra o venta** para el peso colombiano (USDCOP).



Enfoque



Análisis
fundamental



Análisis
técnico

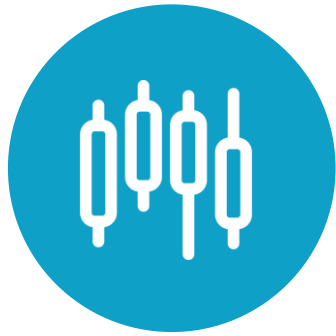


Modelo
Machine Learning



Recomendación

Fuente de datos



Precios de la
divisa

Bloomberg



Información
macroeconómica y
de mercado



Modelo análisis fundamental

Modelo de análisis fundamental

Objetivo: estimar el “precio fundamental” de la divisa.

Datos

77 variables
14 años de historia

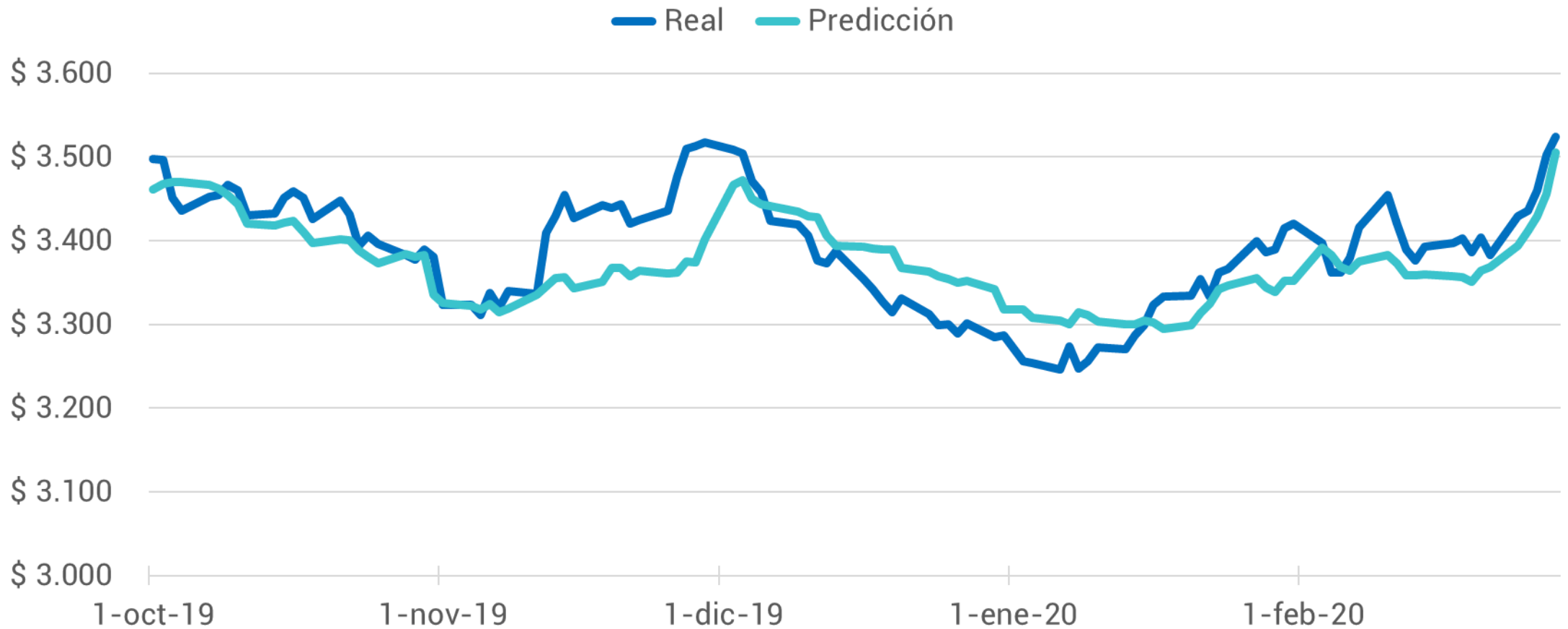
Modelo

Regresión
lineal
56 variables

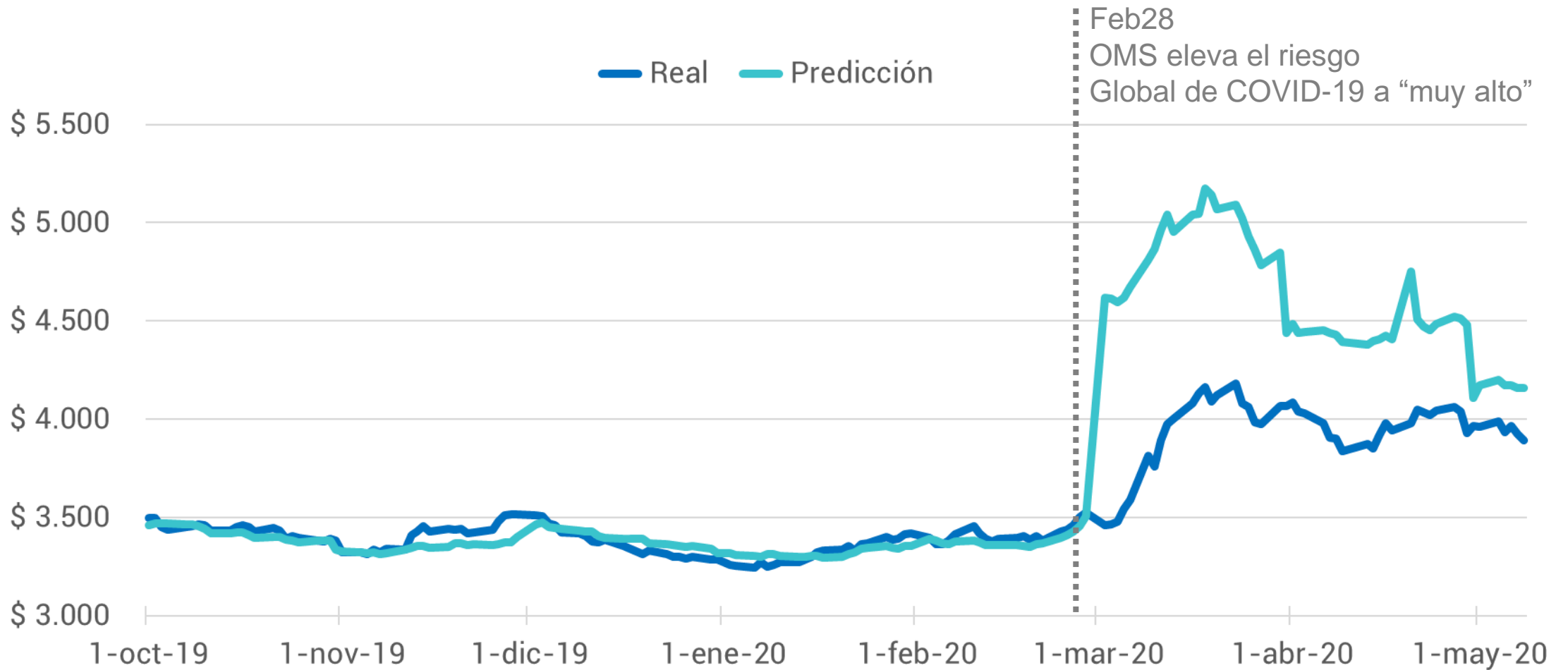
Resultados

Test:
 $R^2 = 0,52$
RMSE = COP\$47,8

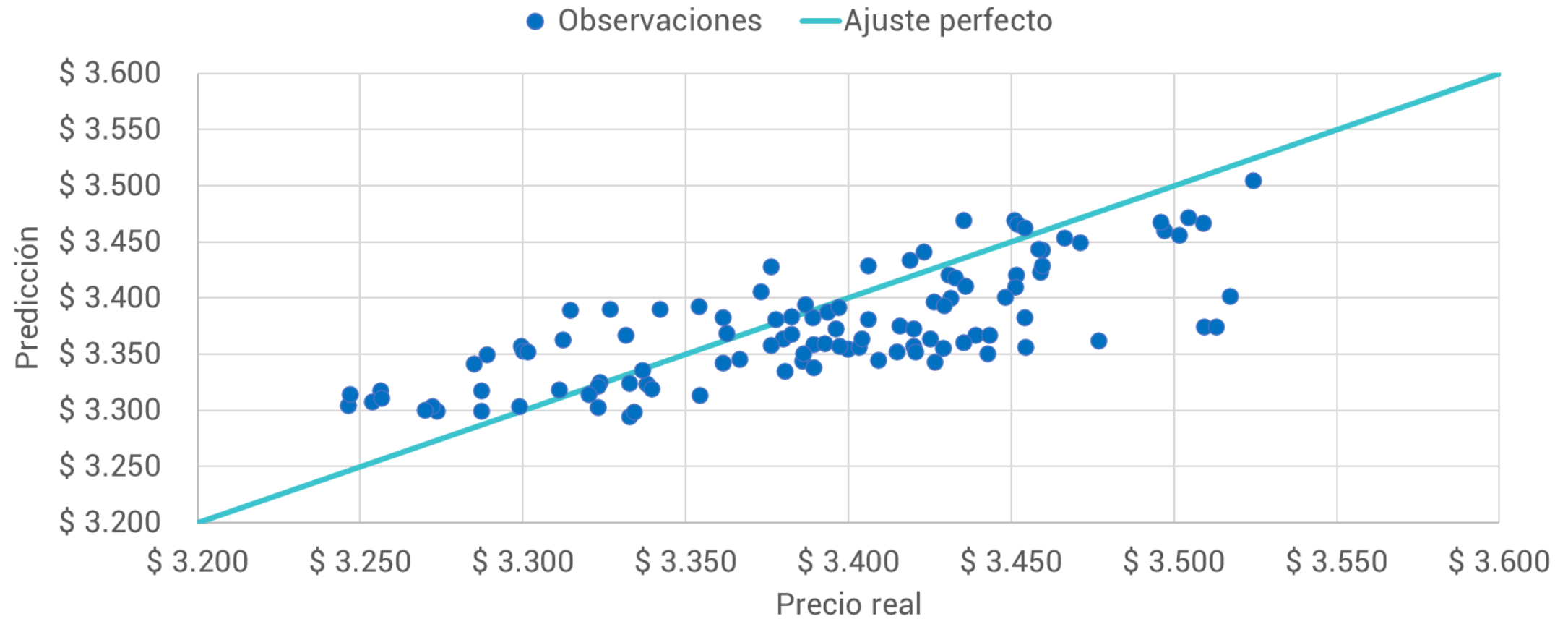
Precio USDCOP (Real vs Predicción)



Precio USDCOP (Real vs Predicción)

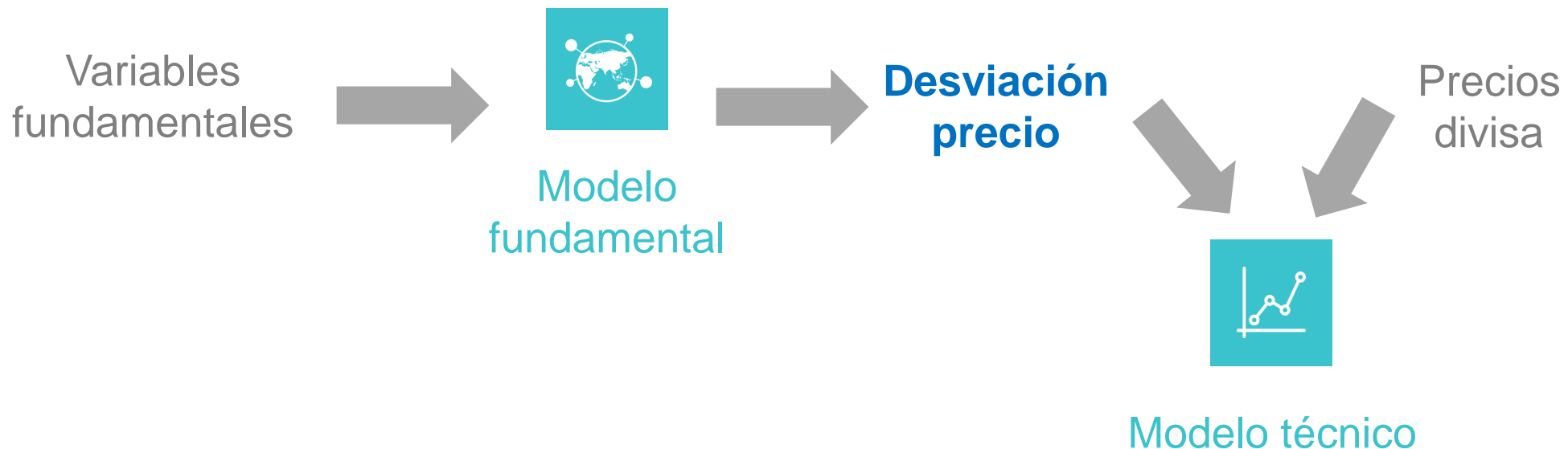


Precio USDCOP (Real vs Predicción)



Integración de análisis fundamental y técnico

La **desviación** del precio fundamental vs el precio de mercado será un **nuevo atributo** que se usará como variable predictora en el modelo de clasificación final.





Modelo análisis técnico

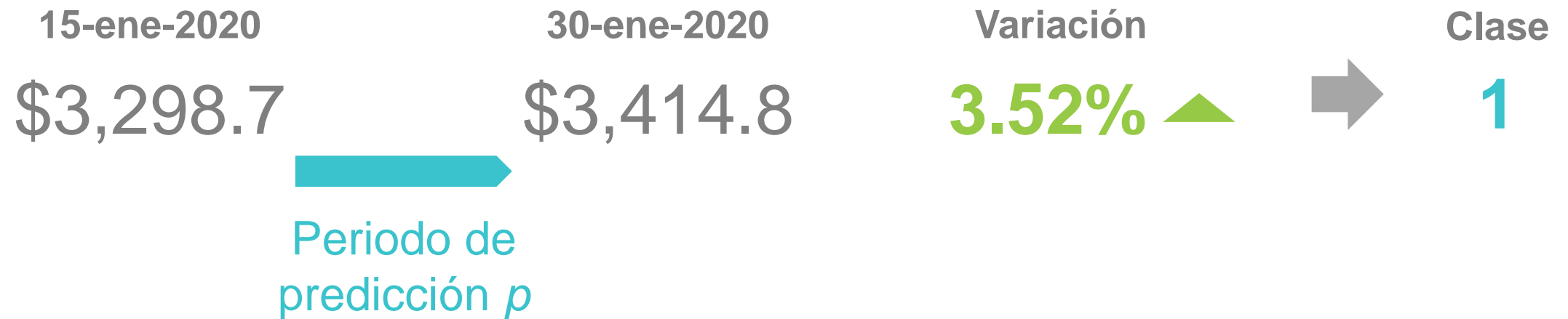
Modelo técnico variables predictoras



Modelo técnico variable de predicción

3 clases $\left\{ \begin{array}{ll} -1 & \text{El precio bajará} \\ 0 & \text{Precio estable} \\ 1 & \text{El precio subirá} \end{array} \right.$ Si $-0.05\% \leq \text{variación} \leq 0.05\%$

Ejemplo

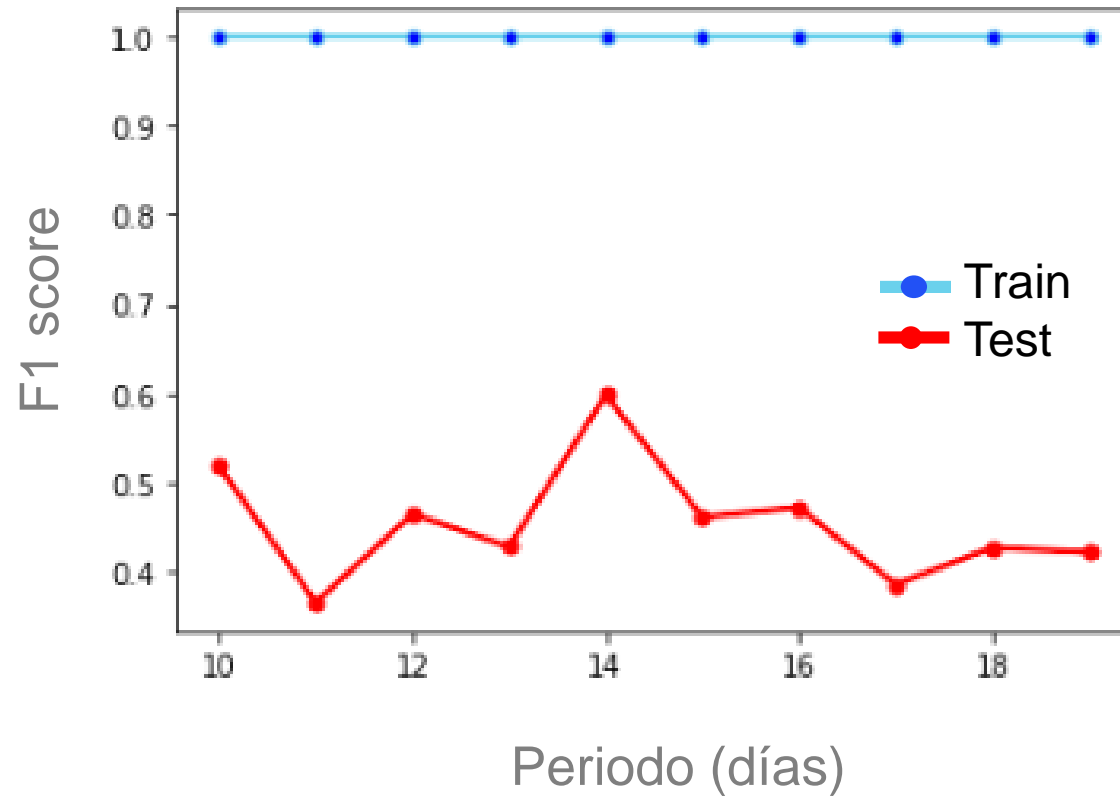


Métrica de evaluación

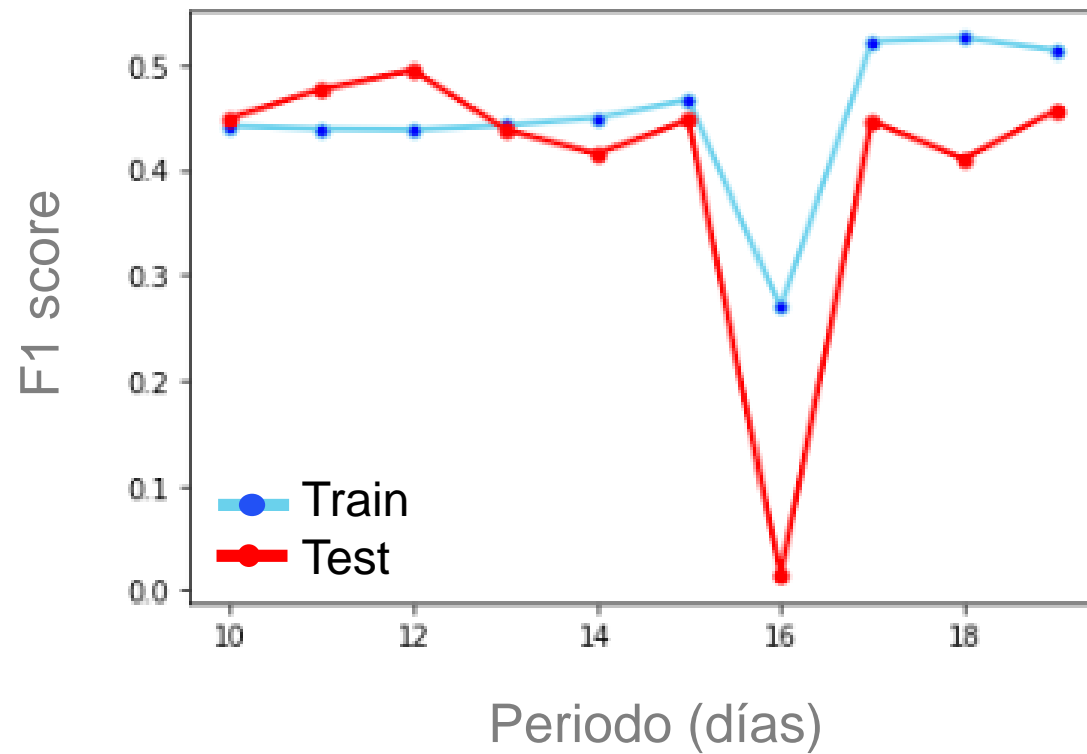


F1 score

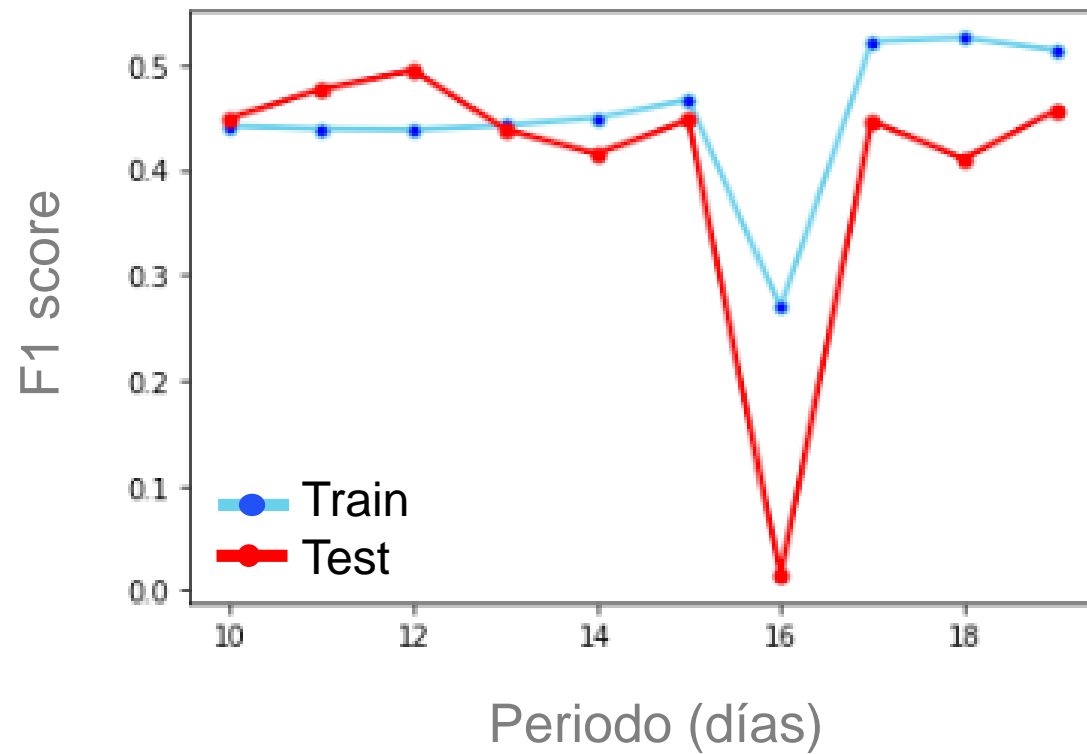
Gradient Boosting



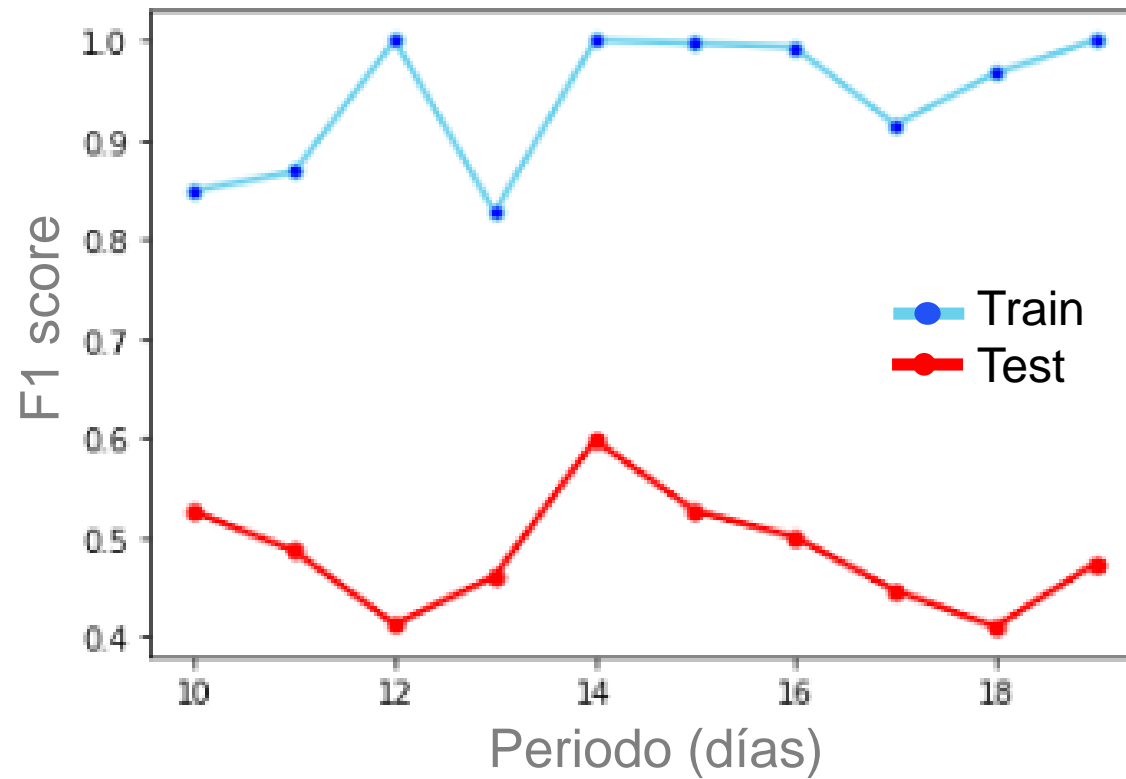
Support Vector Machines



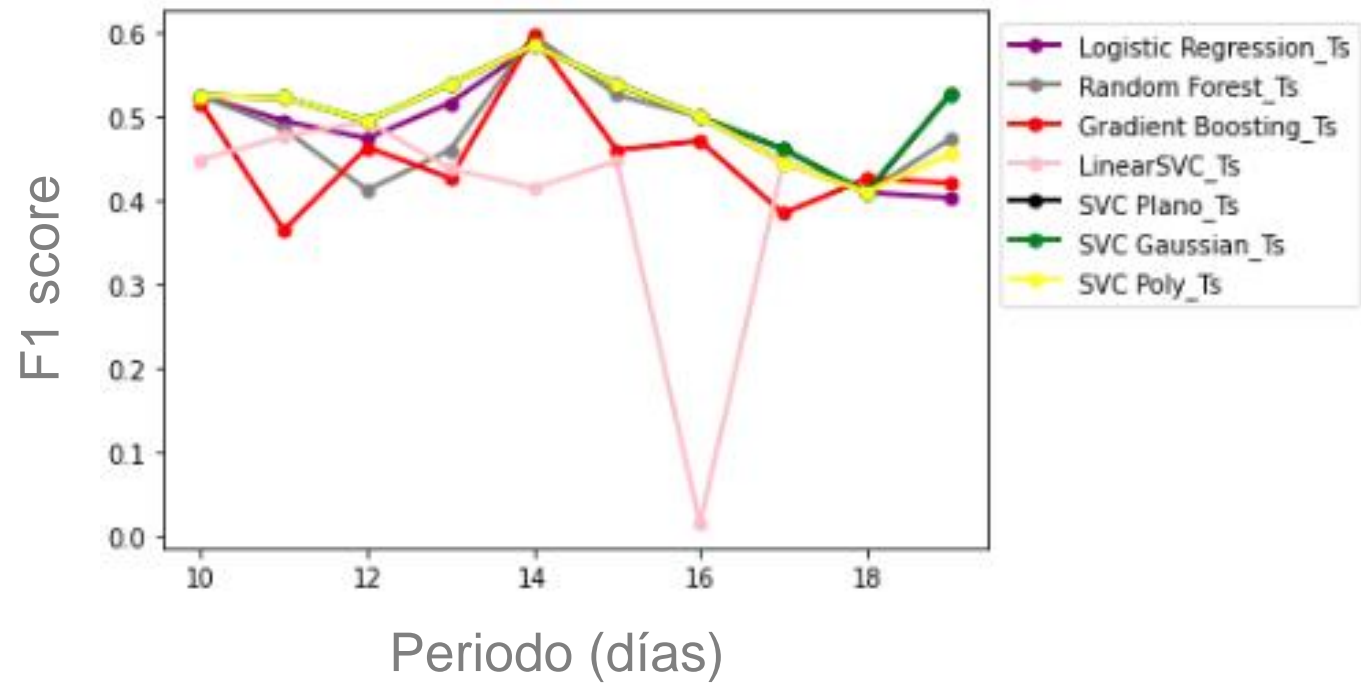
Support Vector Machines



Random Forest



Comparación modelos



Modelo Final



Matriz de confusión

	-1	0	1
-1	11	2	15
0	0	0	1
1	8	0	21

Matriz de desempeño

	Precision	Recall	f1-score
-1	0,58	0,39	0,47
0	0	0	0
1	0,57	0,72	0,64

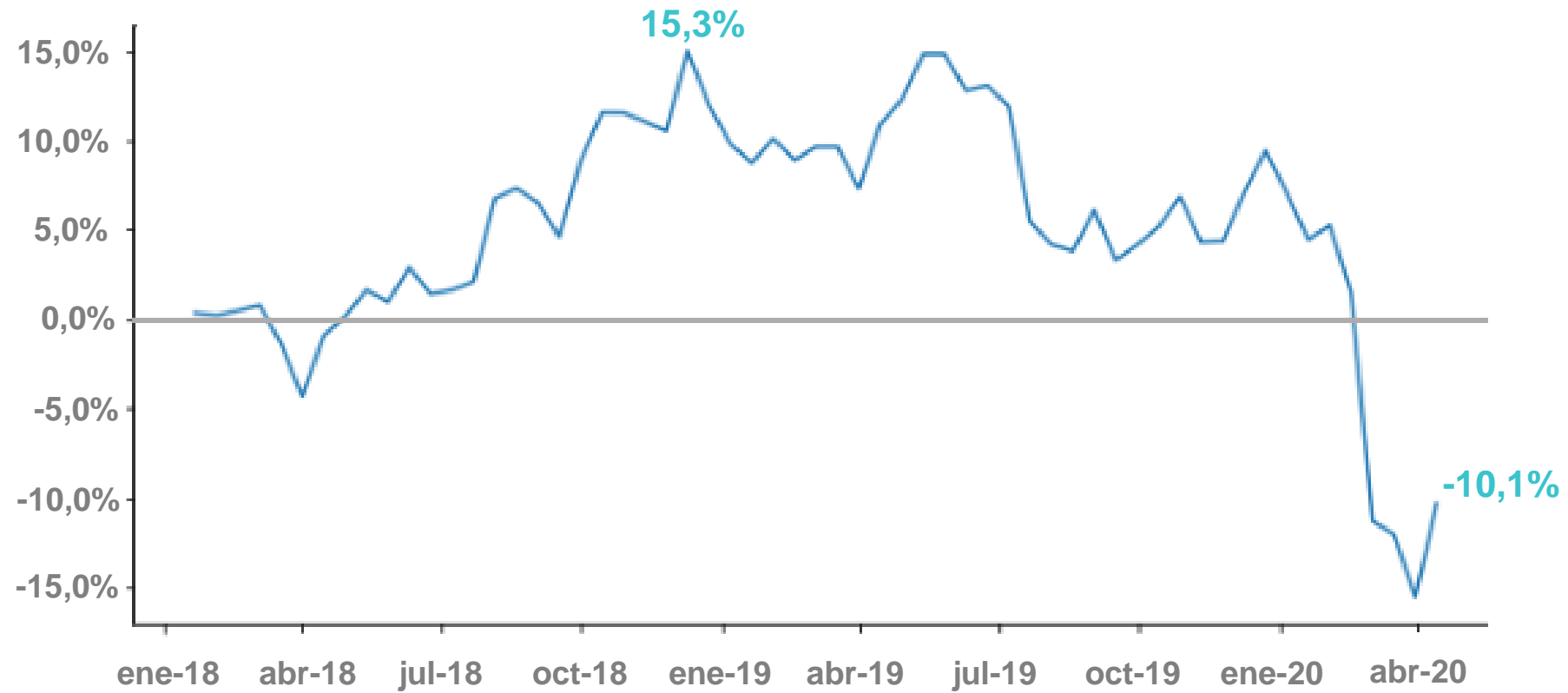


Backtesting

Retorno acumulado de la estrategia USDCOP (Enero 2018 – Febrero 2020)



Retorno acumulado de la estrategia USDCOP (Enero 2018 – Mayo 8 2020)



Conclusiones

- Fue posible desarrollar un modelo de clasificación acertado, el cual recomienda operaciones de compra o venta para periodos de 14 días basado en indicadores financieros técnicos y fundamentales.
- Luego de evaluar varios los modelos, se encontró como mejor modelo el Gradient Boosting con decisiones tomadas cada 14 días para los datos de USDCOP.
- Dado el accuracy cercano al 60% si se utiliza este modelo en un horizonte grande de tiempo se espera tener una buena tasa de aciertos que son el insumo inicial para una estrategia rentable de trading.

Referencias

[1]Huang, J. Z., Huang, W., & Ni, J. (2019). Predicting bitcoin returns using high-dimensional technical indicators. *The Journal of Finance and Data Science*, 5(3), 140-155.

[2]Dash, R., & Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The Journal of Finance and Data Science*, 2(1), 42-57.