**Reporte de hallazgos**

**Informe de Hallazgos**

**Efrén Jiménez**

**1/04/2020**

Contenido

[Introducción 3](#_Toc36675717)

[¿Qué es un patrón de vela? 3](#_Toc36675718)

[Exploración de los datos 4](#_Toc36675719)

[¿Qué atributos en los datos podemos considerar estadísticamente significativos para el problema en cuestión? 5](#_Toc36675720)

[¿Qué información concreta podemos derivar de los datos que tenemos? 6](#_Toc36675721)

[¿Qué métodos probados podemos usar? 7](#_Toc36675722)

[¿Conclusiones? 9](#_Toc36675723)

# Introducción

El tipo de cambio de cada par de dinero puede predecirse utilizando un algoritmo de aprendizaje automático durante el proceso de clasificación. Con la ayuda del modelo de aprendizaje automático supervisado, la tendencia alcista o bajista pronosticada de la tasa de Forex puede ayudar a los comerciantes a tomar la decisión correcta sobre las transacciones de Forex. Los algoritmos de aprendizaje automático en el mercado de intercambio de divisas pueden realizar automáticamente las transacciones de compra / venta.

# ¿Qué es un patrón de vela?

Los patrones de velas proporcionan información sobre la acción del precio en un momento. Si bien los patrones básicos de velas pueden proporcionar una idea de lo que está pensando el mercado, estos patrones más simples a menudo generan señales falsas porque son muy comunes. Todos estos patrones se caracterizan por el precio que se mueve en una dirección, y luego aparecen velas en la dirección opuesta que empujan significativamente a la tendencia anterior.

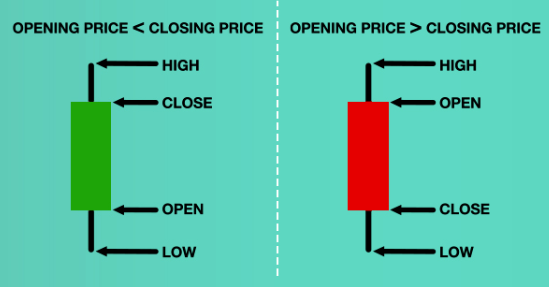


Ilustración Candela de precio

# Exploración de los datos

La cantidad de observaciones analizadas son 10000, del intercambio de monedas de EUR/USD, en estas se puede encontrar con un set de datos balanceado entre candelas de precio arriba y precio abajo.

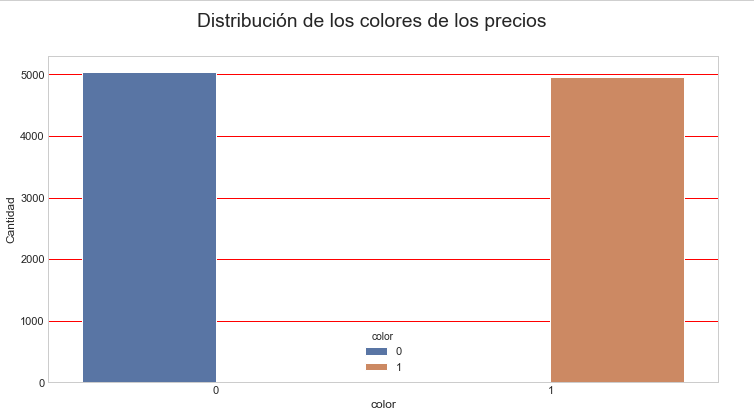


Ilustración Balanceo del set de datos

La cantidad de variables presentes en el set de datos es de 77 columnas.

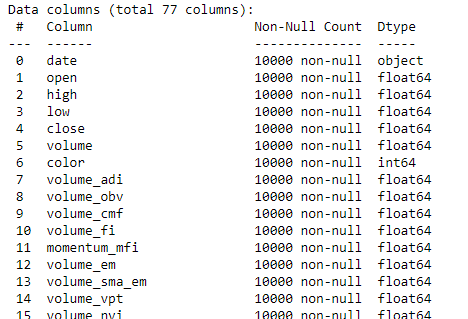


Ilustración Variables y registros

Los rangos de valores del precio de intercambio rondan entre el 1.07 y 1.11 entre las variables: close, high, open, low.

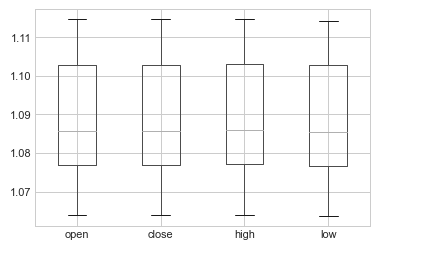


Ilustración Box Plot precios

# ¿Qué atributos en los datos podemos considerar estadísticamente significativos para el problema en cuestión?

Las variables momentum\_rsi, trend\_vortex\_ind\_diff, volatility\_bbp, trend\_cci son las más representativas en modelos de clasificación como SVM, lo cual podrían usarse para realizar algunos análisis sobre estas variables. Se utilizó el modelo PCA para reducir la dimensionalidad del dataset.

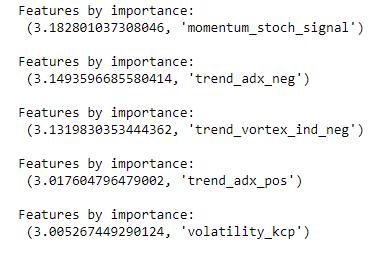


Ilustración Variables luego de realizar el PCA

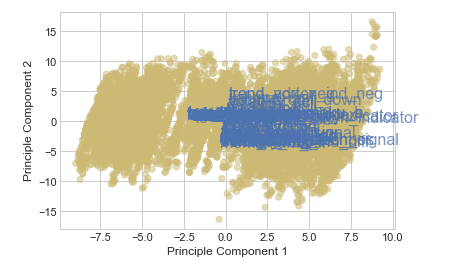


Ilustración Gráfico variables luego de realizar el PCA

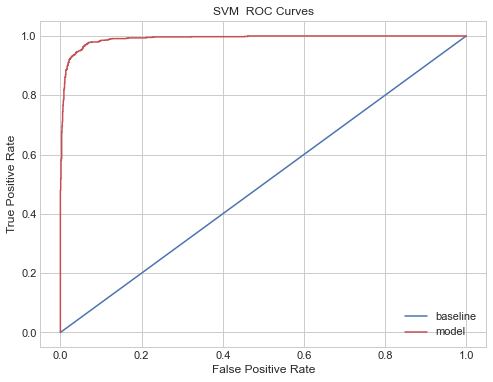
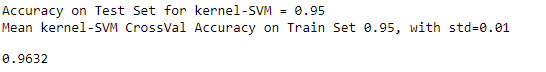
# ¿Qué información concreta podemos derivar de los datos que tenemos?

La exploración de datos del conjunto de información brindado de las tarjetas de crédito muestra las siguientes posibilidades de negocios:

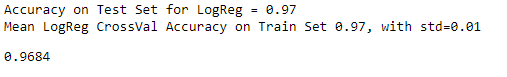
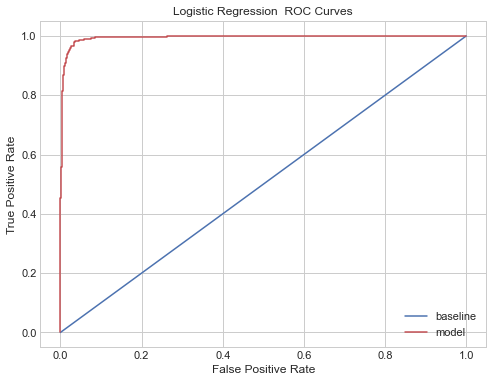
* Los modelos de clasificación pueden ayudar a pronosticar la tendencia de EURUSD hacia arriba o hacia abajo.
* Los problemas de intercambio de divisas ahora pueden verse como problemas de clasificación que pueden resolverse mediante el aprendizaje supervisado.
* El experimento ha demostrado que los modelos de clasificación pueden ayudar a encontrar patrones en datos enormes y complejos para ayudar a los inversionistas con sus decisiones de transacciones.

# ¿Qué métodos probados podemos usar?

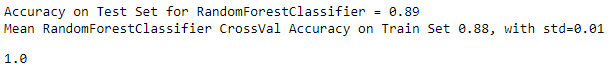
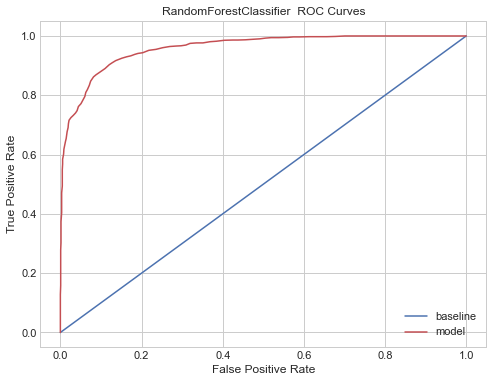
Se verifico el desequilibrio de datos, la visualización de las características y se entendió la relación entre diferentes variables.

 SVM

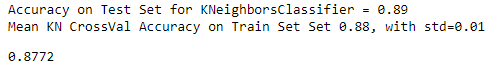
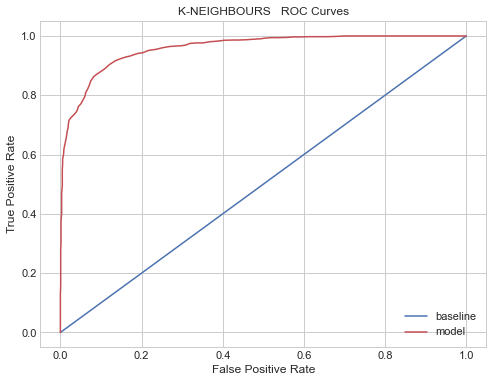
Logistic Regression



Random Forest Classifier



K-NEIGHBOURS



# ¿Conclusiones?

Los modelos SVM,Logistic Regression brindaron resultados alrededor del 96% de probalidad de clasficar correctamente el set de pruebas, el modelo K-NEIGHBOURS alrededor del 87 de probabilidad de clasificar correctamente. El modelo RandomForestClassifier dio un 100% de probabilidad de clasificar correctamente el status del precio, por lo cual no tienen margen de error, por lo cual no se recomienda utilizarlo.

