

Predicción porcentaje de retorno de ganancia.

Leonel Ardila, lfardilap@unal.edu.co
Juan Flórez, jsflorezj@unal.edu.co

May 7, 2018

Abstract

El objetivo de este proyecto es crear un agente inteligente que haga trading con cripto-monedas de forma automática, para ello se implementará una red neuronal Long Short-Term Memory (LSTM) que prediga el porcentaje de variación de valor del Bitcoin frente al Theter (BTC/USDT), la cual será utilizada por el agente como criterio principal para decidir qué tipo de movimientos realizar. Se utilizarán marcos temporales de 1, 5, 15, 30 minutos, 1, 2, 4, 6, 12 horas y por último variaciones en el rango de un día para entrenar la red neuronal LSTM. El agente inteligente tendrá acceso al estado actual del mercado (el valor del precio en los diferentes marcos temporales), con esta información calculará tanto la probabilidad de la variación porcentual del precio en un intervalo de tiempo (e.g entre 1% y 2% % entre 3% y 4%) como el tipo de variación i.e. alcista o bajista, y usará esta información para decidir qué movimiento realizar..

1 Predicción de series de tiempo

Las series de tiempo aparecen naturalmente en el estudio de sistemas dinámicos, como lo son el movimiento de grandes conglomerados de partículas en medios acuosos (movimiento de proteínas) y la variación del precio de una acción en el mercado de valores. Para estudiarlas se han desarrollado varias herramientas estadísticas, e.g. los modelos ARIMA (Auto Regresive Integrated Moving Average).

Se planteó la posibilidad de utilizar modelos de Machine Learning para hacer predicciones de series de tiempo, sin embargo, una aplicación directa de los modelos estándar no mostró ser una opción viable. Tales modelos son eficientes para reconocer patrones y clasificar a partir de estos patrones, y suponen que los datos utilizados representan un ensamble estadístico estático, por ello, terminaban aprendiendo la forma de los datos de entrenamiento y reproduciendola cuando se les pedía realizar una predicción.

Fue necesario introducir nuevos modelos que tuvieran memoria y con ella la capacidad de olvidar conforme predecían más pasos temporales en el futuro, el principal de estos modelos es la red neuronal LSTM (Long Short-Term Memory). Esta es una red neuronal recurrente (RNN) cuyos bloques son del tipo LSTM, i.e. cada bloque tiene una unidad de memoria, una compuerta de entrada, una compuerta de olvido y una compuerta de salida, donde cada compuerta es una neurona. Este modelo suele ser entrenado usando backpropagation, ya que el error se guarda en las unidades de memoria y no es disipado por el tamaño de la red neuronal recurrente, sin embargo, las aplicaciones comerciales de las redes neuronales LSTM utilizan CTC (connectionist temporal classification) como método de entrenamiento.

2 Información para el entrenamiento

Los mercados de criptodivisas mas grandes del mundo (bitfines, bitrex, Binance) cuentan con API's que permiten obtener información histórica del mercado en cada uno de los marcos temporales ya mencionados, información sobre las fluctuaciones hitóricas, el volumen comercializado y las ordenes de compra y venta. Para entrenar la red neuronal se utiliza la información suministrada por la API de Binance.

3 Resultados esperados

Se espera crear un agente inteligente que haga trading con criptodivisas y utilice la pronóstico el porcentaje de variación del mercado provista