ITESO - Maestría en Sistemas Computacionales

Asesor: J. Guadalupe Olascuaga Cabrera

Co asesor: Luis Fernando Gutiérrez Preciado

Alumno: Mawrer Amed Ramirez Martinez

Reporte de Avance de Trabajo de Obtención de Grado

Asesoría: 18 junio de 2020 – IDI4

**Completado:**

**Paper:** FOREX Trend Classification using Machine Learning Techniques

**TIPO:** CLASIFICACION

**DATASET:** USD/YEN, USD/EGP, EURO/EGP and EURO/SAR

The datasets represent a time period of 1852 days from April 2003 to August 2010 excluding the weekends.

**Rago utilizado:**

1-30 dias PRUEBAS

Usados para índices desde 1 dia a 15 dias.

**Modelo: SVM**

**Notas:**

two discriminative feature selection techniques are used, namely and SVM-based technique, and a bagging decision tree technique.

**Paper:** Modeling and Prediction of Time-Series -A Case Study with Forex Data

**TIPO:** Regression

**DATASET:**

USD/JPY

from 2017/12/5 to 2018/10/19 with a resolution of 5 minutes time slot

**Rago utilizado:**

1-30 dias PRUEBAS

Usados para índices desde 1 dia a 15 dias.

**Modelo: SRV and LSTM**

**Notas:**

two discriminative feature selection techniques are used, namely and SVM-based technique, and a bagging decision tree technique.

**Avance:**

Tecnica de análisis usando multiples time Frames:

Trades de días usan todo el dia para monitorear charts, y posteriormente se van a un minuto, 15 min o una hora.

Identifican tendencias principalmente en días y posteriormente sabiendo la tendencia buscan la entrada perfecta.

Si se analiza en Una Hora se busca un entry en 15 min.

Cuando se analiza por días, usan el MA200

En base a artículos leídos referentes a estrategias de trading, mencionan el análisis de forex en escala mayor para la ejecución de entradas o salidas a escala menor.

Es decir, se puede hacer un análisis por días o semanas y en base a esa información pasar a escala de **4H** o **1H** y realizar entradas en forex.

Definitivamente HFT no es algo que busque realizar en este trabajo. Es decir no se buscaría analizar movimientos de segundos y 1 minuto y proceder en esta escala.

Con esto en mente Inicié un análisis de auto correlación de features, sin embargo, lo que encontré fue un ruido extraño en todos los datasets descargados de la plataforma Dukascopy. Compare con otras fuentes y confirme que el Dataset trae un patrón extraño.

Los datos que utilizare serán de: <https://www.histdata.com/download-free-forex-historical-data/?/metatrader/1-minute-bar-quotes/EURUSD>

Hare un append de todos los años y a partir de eso un re sampling para generar los datasets de distintas frecuencias.

**Trabajo para la siguiente sesión:**

* ~~Entender el modelo LSTM (porque 51, porque 25). 51 es el Rango que utilizo, la cantidad de features hacia atrás. 25 es el numero de Features que incluye mi Dataset.~~
* Utilizar el Close también como
* TENER EL INTERVALO que voy a utilizar
  + Diferencia entre gradiente descendiente y gradiente estocástico.
* El **descenso de gradiente estocástico** (**SGD**) lleva esta idea al extremo: usa un solo ejemplo (un tamaño del lote de 1) por iteración. Cuando se dan demasiadas iteraciones, el SGD funciona, pero es muy inconsistente. El término "estocástico" indica que el ejemplo único que compone cada lote se elige al azar. [1]

* + Diferencia entre CELLS y UNITS.
  + Comparar entre intervalos y analizar estadísticamente la volatilidad y definir la mejor. Comparar estrategias y en diferentes intervalos. Algo que sustente la selección de ese intervalo. Euristica debe tener. **likelihood**
* **Utilizar la auto-correlacion para análisis de intervalos**
* Utilizar Cross Validation (porcentajes de pruebas) en modelos.
* **Uso de PCT Change.**
  + Logaritmo de PCT Range e identificar diferencias.
  + Independiente del precio con PCT se puede predecir a donde se mueve.
* **Enfocarme en el análisis del dataset. Y Concluir sobre el análisis.**
* Avanzar mas con los modelos.
  + Implementar lo aprendido en el curso en mis Jupyter Notebooks.
* Probar no solo con el siguiente evento, sino varios adelante. Para LSTM u otros modelos.
* Hacer modelos y combinar Features.
* Se pretende tener una estrategia de trading que pueda tener reglas en base a un modelo o varios. (Ejemplo predecir por minutos, horas y días) en base a los resultados tomar una decisión).

Fuentes:

[1]

“Reducción de la pérdida: Descenso de gradiente estocástico,” *Google Developers*, 2020. [Online]. Available: https://developers.google.com/machine-learning/crash-course/reducing-loss/stochastic-gradient-descent?hl=es-419. [Accessed: 18-Jun-2020].

‌

Notas;

Seguir. Investigando lo del plot con index DateTime

Metricas

AUC para clasificación, derivación del ROC. Reciving operator clasifier

Precision y el Recall

RUC

Matriz de confusión.

Balancear sus datasets. Oversampling QAV, Trades , TBBAV, TBQAV, DownTrade, Uptrade, PctSum, PCT acumulado, MV19, Upper, Lower, Position.

Técnicas de balanceo.