

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FISICOMECAÓNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

PLAN DE TRABAJO DE GRADO

FECHA DE PRESENTACIÓN:

Bucaramanga, 20 de octubre de 2017.

TÍTULO:

Predicción de series financieras con redes neuronales recurrentes

MODALIDAD:

Trabajo de Investigación.

AUTOR:

Edwin Jahir Rueda Rojas
Código: 2131432

DIRECTOR:

Raúl Ramos Pollán

ENTIDAD INTERESADA:

Universidad Industrial de Santander

COMITÉ DE TRABAJOS DE GRADO:

EVALUADOR ASIGNADO:

CONCEPTO DEL EVALUADOR:

APROBACIÓN DEL COMITÉ:

FECHA:

Acta No.

Índice general

Resumen	3
1 PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA	4
2 OBJETIVOS	5
2.1 Objetivo General	5
2.2 Objetivos Específicos	5
3 MARCO DE REFERENCIA	6
3.1 Series Temporales	6
3.2 RNNs.....	6
3.3 Predicción financiera	8
4 METODOLOGÍA	9
4.1 Obtención y Pre-procesado de los datos	9
4.2 Establecer métricas y evaluar el desempeño	9
con métodos clásicos de <i>machine learning</i> .	
4.3 Realizar exploración de arquitecturas RNNs	9
y configuraciones multiseñal.	
4.4 Evaluar el rendimiento de los modelos predictivos	10
y proponer estrategias de <i>trading</i> .	
5 CRONOGRAMA	11
6 PRESUPUESTO	12
7 REFERENCIAS	13

Resumen

Las series de tiempo se encuentran en diferentes áreas, como en las comunicaciones, el transporte, la salud y las finanzas [1]. Las redes neuronales recurrentes (RNN) son un subconjunto de redes neuronales inspiradas en el funcionamiento de una neurona humana debido a su configuración basada en unidades LSTM (*Long Short Term Memory*) las cuales son capaces de recordar características por cierto periodo de tiempo. Estas unidades LSTM han sido utilizadas en diferentes ámbitos tales como la clasificación del espectro de radio solar [2], la búsqueda de contraseñas [3], el aprendizaje de modelos fisiológicos para el comportamiento de la glucosa en la sangre [4], y la predicción en el mercado de valores [5], el cual ha tomado importancia en los últimos años.

Las predicciones sobre las series financieras han sido objeto de investigación durante décadas, pero debido a su complejidad, versatilidad y dinamismo, han resultado ser una tarea complicada. No obstante, a lo largo del tiempo han tomado la iniciativa y han tratado de resolver este problema. Este trabajo de investigación abordará el problema de predicción de series temporales utilizando métodos clásicos de *machine learning*, aunque posteriormente se profundizará en el método específico de RNNs con unidades LSTM ya que como se mencionó anteriormente, estas tienen la capacidad de recordar, siendo esta característica la clave para generar la hipótesis de que las RNNs funcionan mejor en la predicción de series financieras que los métodos clásicos de machine learning [5].

La analítica de datos juega un papel importante a la hora de encaminarnos en el desarrollo de una actividad de predicción, ya que es necesario saber cómo organizar los datos de entrada para su posterior procesamiento y así obtener los datos de salida esperados. Para poder entender los datos de series financieras, se tienen que tener en cuenta tres palabras claves, la primera es el *Bid* que es el precio al que el mercado está dispuesto a comprar, o sea al precio que uno puede vender, la segunda es el *Ask* que es el precio al cual el mercado está dispuesto a vender, o sea al precio que uno puede comprar, y la tercera es el *Spread* el cual es la diferencia entre el *Bid* y el *Ask*.

Palabras clave: Machine learning, analítica de datos, redes neuronales recurrentes, series de tiempo, bid, ask, spread, LSTM.

1 PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA

Las redes neuronales recurrentes (RNNs), son una clase de redes neuronales artificiales que se han aplicado exitosamente en varios entornos tales como el reconocimiento de voz [6], la generación de texto [7], la generación de geometría algebraica (Latex) [8], la traducción de video a frases [9], la generación de títulos de imágenes [10], la verificación de usuarios basada en la dinámica de pulsaciones de teclas [11], entre otros. Especialmente, las RNNs han mostrado su eficiencia para la predicción de series temporales a través de unidades LSTM (*Long Short Term Memory*), ya que estas tienen la capacidad de actuar como una “memoria” permitiendo a la RNN recordar lo que necesita y olvidar lo que no es necesario.

En el ámbito financiero el uso de métodos de *machine learning* es relativamente reciente y el uso de las RNNs está empezando a ser explorado ya que se intuye que las señales financieras tienen una cierta clase de patrones periódicos [12]. Por esto, se espera que las RNNs sean capaces de predecir características futuras de las señales en base al contenido informativo de lo aprendido.

Las RNNs están diseñadas para aceptar como entrada múltiples señales para posteriormente predecir una sola. En el ámbito de las finanzas, el uso de múltiples señales se refiere a distintos valores financieros que pueden influir entre ellos (p.ej. el valor bursátil de distintas compañías del mismo sector, la tasa de cambio de monedas de economías relacionadas, etc. Si nos basamos en datos históricos, podemos asumir que las distintas señales financieras son interdependientes y por tanto la información sobre unas se puede usar para predecir instrumentos de otras.

Por tanto, el principal objetivo de este trabajo de investigación es explorar el uso de RNNs para predecir señales financieras combinando múltiples entradas de otras señales e integrar dichas predicciones en estrategias de *trading*.

2 OBJETIVOS

2.1 OBJETIVO GENERAL

Diseñar y evaluar redes neuronales recurrentes para la predicción de señales financieras basadas en múltiples señales.

2.2 OBJETIVOS ESPECÍFICOS

- Adquirir datos financieros de diferentes fuentes.
- Desarrollar un pre-procesado de los datos multiseñal obtenidos anteriormente.
- Establecer métricas de predicción y tiempo de ejecución para evaluar el rendimiento del objetivo.
- Establecer una línea base de desempeño predictivo con métodos clásicos de *machine learning*.
- Realizar una exploración de arquitecturas RNNs y configuraciones multiseñal.
- Evaluar el rendimiento de los modelos predictivos y proponer estrategias de *trading*.

3 MARCO DE REFERENCIA

3.1 Series Temporales

La predicción del movimiento de los índices bursátiles del mercado es un problema de aprendizaje asociado a las series temporales, el cual se estudia con frecuencia ya que es una tarea desafiante debido a la volatilidad y los problemas de no linealidad que presentan [13]. Aunque estudios previos se basaban en algoritmos no recurrentes, en las últimas décadas las redes neuronales recurrentes cada vez van tomando más importancia en este ámbito [14], ya que estas tienen la característica de poder recordar ítems a lo largo del tiempo.

La importancia de predecir una serie de tiempo es esencial debido a que siempre que se realiza una planificación en algún negocio o proyecto, este debe partir de un conocimiento de incertidumbre sobre las condiciones futuras. Por ejemplo, en el mercado de valores es imprescindible tener cierta confianza de que el mercado va tender a la alta o a la baja para así poder efectuar transacciones favorables. Esta confianza durante décadas la han brindado diferentes herramientas de predicción, tales como los modelos lineales de ARMA, ARIMA, filtros de Kalman, etc. Y modelos no lineales como las redes neuronales, estimadores bayesianos, árboles de regresión, entre otros [15].

Una serie temporal es una colección de observaciones de una variable adquiridas secuencialmente en el tiempo [16], en la *figura 1* se muestra el comportamiento del Euro en relación con el Dólar, la gráfica nos muestra como el *Bid* cambia a lo largo del tiempo, dándonos una muestra de la bursatilidad de dichas series temporales.

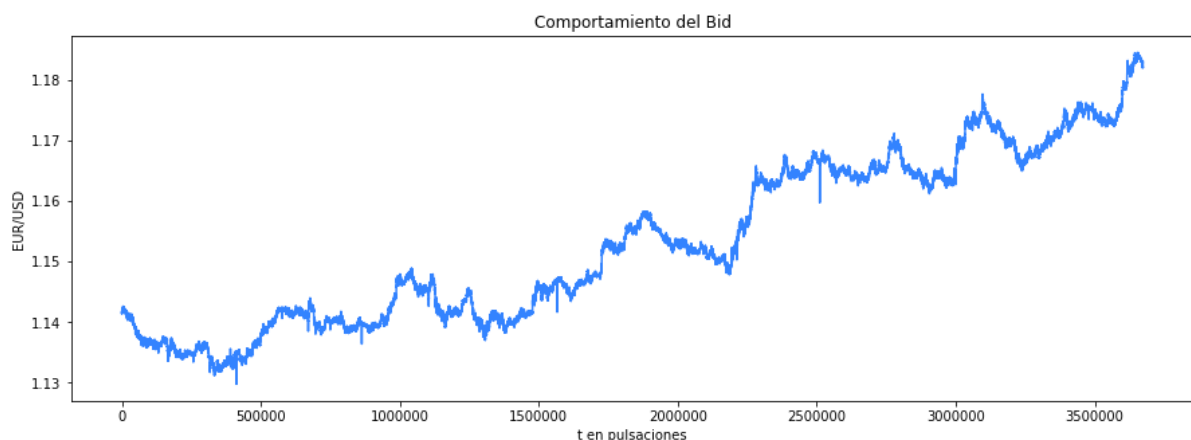


Figura 1. Representación del Bid, datos tomados de TrueFX

3.2 Recurrent Neural Network (RNN)

Las unidades LSTM las cuales son utilizadas frecuentemente cuando se trata de predecir series temporales, son un conjunto de subredes conectadas recurrentemente, lo cual les permite solucionar el problema conocido como *vanishing error problem*, que consiste en como la influencia de las entradas pasadas va decayendo rápidamente en el tiempo, puesto que para series financieras, es imprescindible que el modelo con el que se trabaje tenga la capacidad de recordar características pasadas para tener cierta tendencia del comportamiento de la señal. Este problema, las LSTM lo resuelven mediante el uso de las celdas de memoria [15], permitiéndole así, recordar características que se consideran relevantes para su posterior uso.

El funcionamiento de una RNN está dado por unas unidades de entrada $X(t)$ en un cierto instante de tiempo, estas a su vez pasan a unas capas ocultas que llamaremos 'A' como se ilustra en la *figura 2*, en las cuales aparecen las celdas de memoria LSTM, y tiene una capa de salida $h(t)$ la cual entrará en el siguiente instante de tiempo si así lo requiere la red. El número de capas ocultas depende del ajuste del diseñador de la RNN, estas varían en concordancia a las necesidades de las tareas [15].

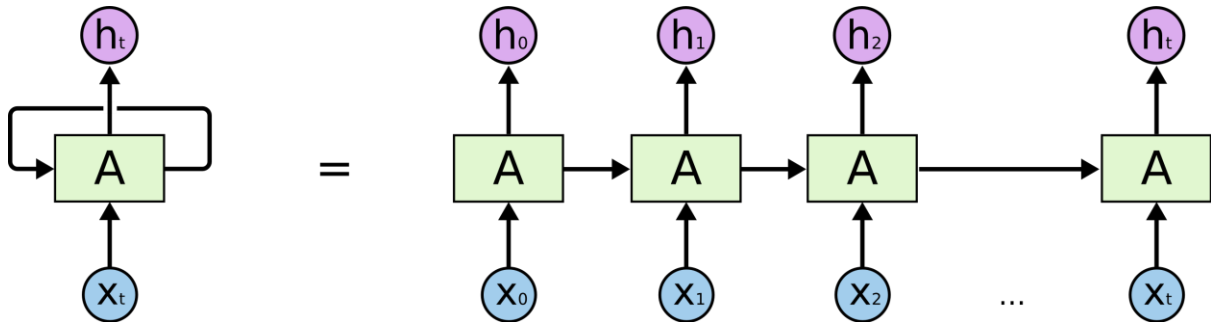


Figura 2. Representación de una RNN, tomado de colah's blog: Understanding LSTM networks.

Un bloque de unidad LSTM contiene unos parámetros de entrada, los cuales intervienen con ciertas operaciones matemáticas que son reflejadas en la *figura 3*, estos a su vez producen una salida esperada y un bloque de memoria para el siguiente instante de tiempo el cual se activará o no, según lo considere la red.

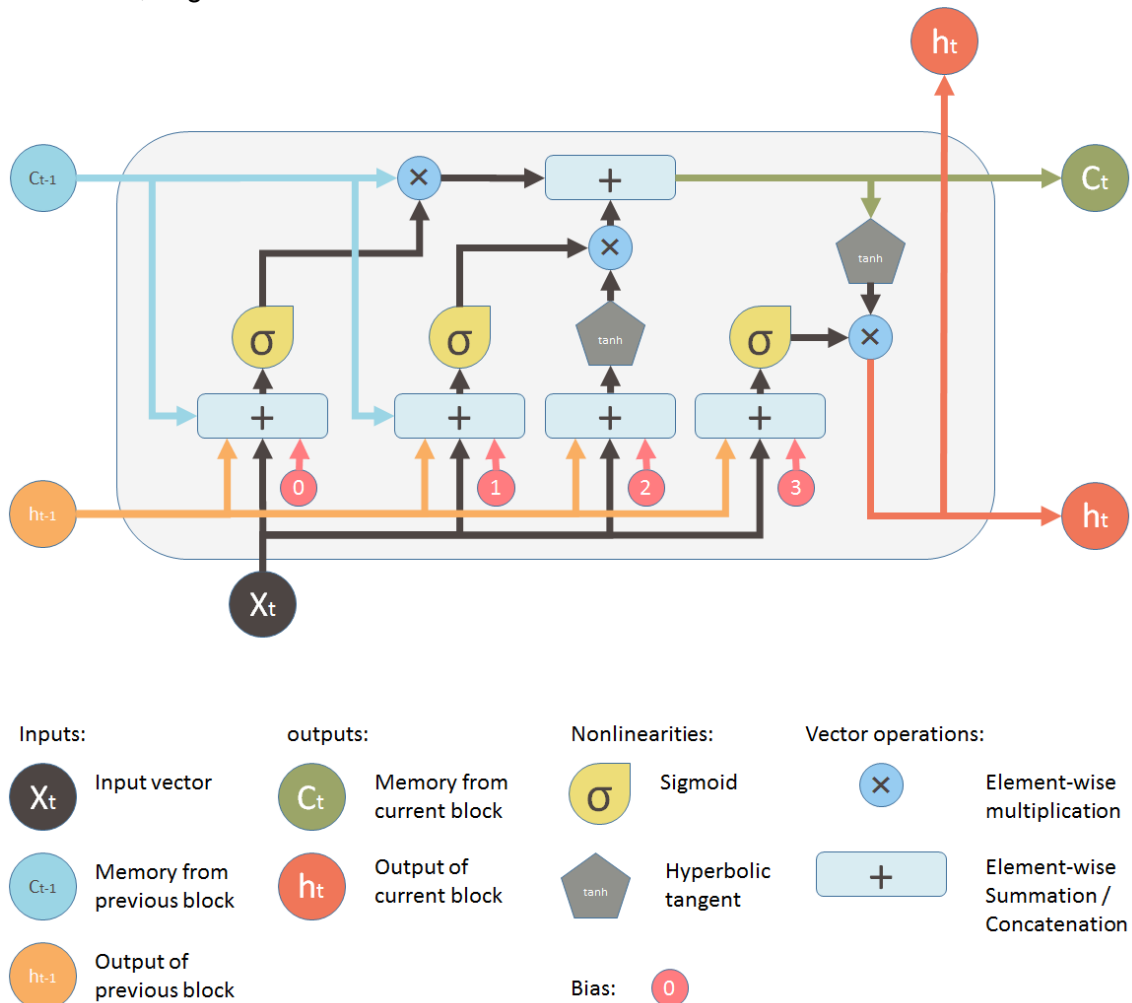


Figura 3. Unidad LSTM, tomado de Shi Yan: Understanding LSTM and its diagrams.

3.3 Predicción financiera

En la actualidad la predicción financiera juega un papel importante en la toma de decisiones futuras, tales como en la compra y venta de acciones, ya que es muy importante tener cierta confianza de cómo se comportará un instrumento financiero para poder invertir en él. Estudios recientes han demostrado que estas predicciones tienden a dar confianza en cuanto al mercado financiero se refiere, pues los estudios han arrojado resultados favorables [14].

Aunque las RNNs no están arraigadas completamente a la predicción financiera, han tomado mayor importancia debido a que existe la hipótesis de que pueden funcionar mejor que los modelos planteados anteriormente para esta tarea debido a su capacidad de recordar valores para su posterior uso. [12].

4 METODOLOGÍA

Para ejecutar el trabajo de investigación se propone un desarrollo metodológico el cual abarca un conjunto de pasos que van desde la obtención de datos financieros de distintas fuentes hasta evaluar el modelo predictivo y proponer estrategias de *trading*. A continuación, se enumeran cada una de las etapas planeadas para el desarrollo del trabajo de investigación, así como las actividades a realizar en cada etapa.

4.1. Obtención y Pre-procesado de los datos

En esta etapa inicial se obtendrán conjuntos de datos de series financieras de distintas fuentes a los cuales se les hará su respectivo pre-procesado para su posterior uso. A continuación, se enumeran las actividades a realizar en esta etapa:

- Obtención de datos financieros de fuentes variables.
- Pre-procesado de los datos según corresponda.

4.2. Establecer métricas y evaluar el desempeño con métodos clásicos de *machine learning*

En esta etapa se establecerán las métricas con las cuales posteriormente se evaluará el desempeño obtenido. Para ello se tratarán los datos pre-procesados obtenidos en la etapa 1 para su posterior uso al entrenar modelos clásicos de *machine learning*, para luego ver el rendimiento de dichos modelos. A continuación, se enumeran las actividades a realizar en esta etapa:

- Establecer las métricas con las cuales se evaluará el desempeño.
- Utilizar métodos clásicos de *machine learning* para la predicción de series temporales utilizando múltiples configuraciones.
- Reportar los resultados obtenidos del proceso de aprendizaje.

4.3. Realizar exploración de arquitecturas RNNs y configuraciones multiseñal

En esta etapa se explorarán diferentes tipos de arquitecturas de RNNs para posteriormente sacar una estimación de cuál sería la configuración de la RNN más apropiada para la predicción de las señales financieras. Por otro lado, también se observarán diferentes configuraciones multiseñal. A continuación, se enumeran las actividades a realizar en esta etapa:

- Explorar diferentes tipos de arquitecturas RNNs.
- Elegir la arquitectura que se acople mejor para la predicción de señales financieras.
- Explorar las diferentes configuraciones multiseñal, para así escoger la acorde.

4.4. Evaluar el rendimiento de los modelos predictivos y proponer estrategias de *trading*

En la etapa final lo que se querrá evaluar es el rendimiento de los modelos predictivos propuestos en etapas anteriores, y basándonos en dichos resultados proponer estrategias de *trading*. A continuación, se enumeran las actividades a realizar en esta etapa:

- Evaluar el rendimiento de los modelos predictivos propuestos anteriormente.
- Basados en estos resultados proponer estrategias de *trading*.

5 CRONOGRAMA

Para poder realizar este trabajo de investigación, se tendrán en cuenta las 16 semanas con las cuales cuenta el semestre, y se hará un cronograma evidenciando lo que se ejecutará por cada semana, teniendo en cuenta que semanalmente habrá reunión con el director del proyecto para aclarar dudas y seguir avanzando. A continuación, se muestra el cronograma:

Actividades	Meses / Semanas															
	1				2				3				4			
Obtención y pre-procesado de los datos																
- Obtención de datos financieros de fuentes variables																
- Pre-procesado de los datos según corresponda																
Establecer métricas y evaluar el desempeño con métodos clásicos de <i>machine learning</i>																
- Establecer métricas con las cuales se evaluará el desempeño																
- Utilizar métodos clásicos de <i>machine learning</i>																
- Entrenamiento de los modelos																
- Evaluar el desempeño de dichos modelos																
Realizar exploración de arquitecturas RNNs y configuraciones multiseñal																
- Explorar diferentes tipos arquitecturas RNNs																
- Elegir la arquitectura más acorde																
- Entrenar modelos RNNs																
Evaluar el rendimiento de los modelos predictivos y proponer estrategias de <i>trading</i>																
- Evaluar modelos RNNs																
- Proponer estrategias de <i>trading</i>																
- Elaborar el documento final																

Tabla 1. Cronograma de actividades

6 PRESUPUESTO

Para la realización de este proyecto de investigación se tiene el siguiente presupuesto, que incluye todos los gastos que se realicen de recursos humanos, equipos tecnológicos, y demás materiales ofimáticos que se requieran. En la siguiente tabla se hace la descripción del presupuesto:

Función	Hora / semana	Valor hora	Tiempo total	Valor total
Director de proyecto	2	\$200.000	32	\$6'400.000
Autor del proyecto	28	\$10.000	448	\$4'480.000
			Total	\$10'880.000

Tabla 2. Presupuesto de recursos humanos

Descripción	Cantidad	Valor unidad	Valor total
Computador	1	\$2'200.000	\$2'200.000
Papelería	-	\$100	\$20.000
Cuaderno	1	\$2.000	\$2.000
Materiales variados	-	-	\$20.000
		Total	\$2'242.000

Tabla 3. Presupuesto de recursos tecnológicos y materiales variados.

Descripción	Valor
Recursos humanos	\$10'880.000
Recursos tecnológicos y materiales variados	\$2'242.000
Total	\$13'122.000

Tabla 4. Presupuesto total.

Observaciones:

- Los gastos de recursos humanos que están reflejados en este presupuesto son asumidos por el autor y director del proyecto.
- Los gastos de recursos tecnológicos y materiales ofimáticos son asumidos por el autor del proyecto.

7 REFERENCIAS

- [1] Sharat C Prasad, Member, IEEE, and Piyush Prasad, *Deep Recurrent Neural Networks for TimeSeries Prediction*.
- [2] Xuexin Yu, Long Xu, Lin Ma, Zhuo Chen, Yihua Yan, *Solar Radio Spectrum Classification with LSTM*.
- [3] Lingzhi Xu, Can Ge, Weidong Qiu, Zheng Huang, Jie Guo, Huijuan Lian, *Password Guessing Based on LSTM Recurrent Neural Networks*.
- [4] Ben Athiwaratkun and Jack W. Stokes, *Malware Classification with LSTM and Gru Language Models and a Character-Level CNN*.
- [5] David M. Q. Nelson, Adriano C. M. Pereira, Renato A. de Oliveira, *Stock Market's Price Movement Prediction With LSTM Neural Networks*.
- [6] Alex Graves, Abdel-rahman Mohamed and Geoffrey Hinton, *Speech recognition with deep recurrent neural networks*. Department of Computer Science, University of Toronto.
- [7] Ilya Sutskever, James Martens, Geoffrey Hinton, *Generating Text With Recurrent Neural Networks*. University of Toronto.
- [8] Andrej Karpathy, *The Unreasonable Effectiveness of Recurrent Neural Networks*. May 21, 2015.
- [9] Haonan Yu, Jiang Wang, Zhiheng Huang, Yi Yang, Wei Xu, *Video Paragraph Captioning Using Hierarchical Recurrent Neural Networks*. Purdue University, Baidu Research - Institute of Deep Learning.
- [10] Christine Donnelly, *Image Generation with Recursive Neural Networks*. Stanford University.
- [11] Pawel Koboжек and Khalid Saeed, *Application of Recurrent Neural Networks for User Verification based on Keystroke Dynamics*. Warsaw University of Technology, Warsaw Poland.
- [12] Kyle Polich, *Recurrent Neural Networks and the Stock Market*.
- [13] Mohammad Sultan Mahmud, Phayung Meesad, *Time Series Stock Price Prediction using Recurrent Error based Neuro-Fuzzy System with Momentum*.
- [14] Jason Poulos, *Predicting Stock Market Movement with Deep RNNs*.
- [15] Pilar Gómez-Gil, *On the use of 'Long-Short Term Memory' neural networks for time series prediction*.
- [16] Jm Marín, *Series Temporales*.

Figura 1, tomada de www.truefx.com, Historical Tick-by-Tick Data.

Figura 2, tomada de “<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>”, colah’s blog: Understanding LSTM networks.

Figura 3, tomada de “<https://medium.com/@shiyang/understanding-lstm-and-its-diagrams-37e2f46f1714>”, Understanding LSTM and its diagrams.