

Jaasiel Logan dos Santos de Carvalho

**Predição da Tendência de Séries Temporais do
Mercado**

**FOREX Utilizando Indicador Baseado em
Estatística Descritiva e Redes Neurais
Recorrentes**

Salvador

2017

Jaasiel Logan dos Santos de Carvalho

**Predição da Tendência de Séries Temporais do Mercado
FOREX Utilizando Indicador Baseado em
Estatística Descritiva e Redes Neurais Recorrentes**

Monografia apresentada ao Colegiado de Sistemas de Informação da Universidade do Estado da Bahia como parte integrante dos requisitos necessários para obtenção do diploma de Bacharel em Sistemas da Informação.

Universidade do Estado da Bahia – UNEB

Departamento de Ciências Exatas e da Terra I – DCET I

Colegiado de Sistemas de Informação

Orientador: Prof. Dr. Leandro Santos Coelho de Souza

Coorientador:

Salvador

2017

Jaasiel Logan dos Santos de Carvalho
Predição da Tendência de Séries Temporais do Mercado
FOREX Utilizando Indicador Baseado em
Estatística Descritiva e Redes Neurais Recorrentes/ Jaasiel Logan dos Santos de
Carvalho. – Salvador, 2017-
54 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Leandro Santos Coelho de Souza

Monografia (Graduação) – Universidade do Estado da Bahia – UNEB
Departamento de Ciências Exatas e da Terra I – DCET I
Colegiado de Sistemas de Informação, 2017.

1. FOREX. 2. Rede Neural Recorrente. 3. Indicador 4. Predição I. Leandro
Santos Coelho de Souza

CDU 02:141:005.7

Jaasiel Logan dos Santos de Carvalho

Predição da Tendência de Séries Temporais do Mercado FOREX Utilizando Indicador Baseado em Estatística Descritiva e Redes Neurais Recorrentes

Monografia apresentada ao Colegiado de Sistemas de Informação da Universidade do Estado da Bahia como parte integrante dos requisitos necessários para obtenção do diploma de Bacharel em Sistemas da Informação.

Trabalho aprovado. Salvador, xx de junho de 2017:

**Prof. Dr. Leandro Santos Coelho de
Souza**
Orientador

Professor
Convidado 1

Professor
Convidado 2

Salvador
2017

*Este trabalho é dedicado à minha mãe, Ana Cristina.
À minha irmã, Jade e a minha companheira Michele.*

Agradecimentos

Aos meus pais, em especial a minha mãe que incansavelmente trabalhou em todos os sentidos para me educar em meio a tantas dificuldades.

A minha irmã que sempre esteve comigo me apoiando.

A minha companheira que pacientemente tanto me escutou e ajudou nos momentos de confusão.

A todos os professores que contribuíram para minha formação sem pedir nada em troca além de dedicação.

E a todos aqueles que me ajudaram e acreditaram no meu potencial durante a minha jornada acadêmica até aqui.

*“Any sufficiently advanced technology
is indistinguishable from magic.”
(Arthur C. Clarke)*

Resumo

Realizar previsões de séries temporais por meio do uso da inteligência artificial (IA) e seus modelos adaptativos vem se tornando regra entre especialistas da área financeira. Assim, espera-se que a IA, através das redes neurais artificiais e seus variados tipos de arquitetura possam auxiliar a operação em um dos mercados mais voláteis existente, o mercado FOREX. Acredita-se que o movimento desse mercado se repita de tempos em tempos e para prevê-lo, entusiastas e especialistas recorrem a métodos de análise dos dados gerados por este mercado, como a análise técnica. Nesse sentido, este trabalho teve o objetivo de propor um indicador que fosse aplicável para predição da tendência de séries temporais mediante análise técnica, a fim de auxiliar a tomada de decisão em relação as operações nesse mercado. Para alcançar esse objetivo, foi desenvolvido um modelo de rede neural, que, quando empregado em conjunto do indicador, resultou em taxas de acerto consideráveis, tendo em vista a complexidade do problema. Os resultados alcançados indicam que esses modelos tem potencial para melhorias e que seu estudo e desenvolvimento podem trazer ganhos financeiros.

Palavras-chaves: FOREX. Redes Neurais Recorrentes. Indicador. Predição.

Abstract

Time series predictions through artificial intelligence (AI) and its adaptive models has become a rule among financial experts. Thus, it is expected that AI through artificial neural networks and their various types of architecture, can help the operation in one of the most volatile markets, the FOREX market. It is believed that the movement of this market is repeated from time to time and in order to predict it, enthusiasts and experts resort to methods of analyzing the data generated by this market such as technical analysis. In this sense, this work had the objective of proposing an indicator that would be applicable to predict the trend of time series by means of technical analysis, in order to help decision making in relation to the operations in that market. To achieve this goal, a neural network model was developed, which, when used in conjunction with the indicator, resulted in considerable success rates, given the complexity of the problem. The results indicate that these models have potential for improvement and that their study and development can bring financial gains.

Key-words: FOREX. Recurrent Neural Networks. Indicator. Prediction.

Lista de ilustrações

Figura 1 – Série temporal do mercado FOREX retirada da ferramenta MetaTrader - EXNESS referente ao par de moeda EUR/USD no período de 19 a 28 de dezembro de 2016	18
Figura 2 – Série temporal genérica	19
Figura 3 – Modelo de Neurônio Artificial (Adaptado de: Haykin (2001, p. 36))	26
Figura 4 – Arquitetura básica de uma RNN e seu loop (Olah (2015))	28
Figura 5 – Representação da arquitetura RNN com seu loop expandido (Olah (2015))	28
Figura 6 – Representação interna de um módulo que compõe a arquitetura LSTM Olah (2015))	29
Figura 7 – Modelo de previsão proposto por (Bortoli (2015))	31
Figura 8 – Modelo de previsão	32
Figura 9 – Fluxo metodológico	33
Figura 10 – Tabela com todos valores calculados	36
Figura 11 – Série do mercado FOREX referente ao dia analisado	37
Figura 12 – Médias referentes ao dia analisado	37
Figura 13 – Divisão dos dados para validação cruzada	41
Figura 14 – Resultados preliminares dos testes das redes LSTM. 30 e 40 épocas	45
Figura 15 – Resultados preliminares dos testes das redes LSTM. 50 e 60 épocas	45
Figura 16 – Valores preditos e esperados.	47
Figura 17 – Arquitetura LSTM	54

Lista de tabelas

Tabela 1 – Hiper-parâmetros da rede	40
Tabela 2 – Novo conjunto de dados criado	40
Tabela 3 – Valores referência para transformação	43
Tabela 4 – Arquiteturas utilizadas para treinamento	45
Tabela 5 – Arquitetura de rede com melhor desempenho	46
Tabela 6 – Taxa de acerto da rede sobre o conjunto de testes	46

Lista de abreviaturas e siglas

ARIMA	Auto-Regressivo Integrado de Média Móvel
ARCH	<i>Autoregressive Conditional Heterocedasticity</i>
EUR	Euro
FOREX	Foreign Exchange
FTDNN	Rede Neural alimentada Adiante Focada Atrasa no Tempo
IA	Inteligência Artificial
LSTM	<i>Long Short Term Memory</i>
MQL4	<i>MetaQuotes4</i>
MSE	Erro Quadrático Médio
MT4	<i>MetaTrader4</i>
PIP	<i>Percentage In Point</i>
RNA	Rede Neural Artificial
RN	Rede Neural
RNR	Rede Neural Recorrente
SGD	Gradiente Descendente Estocástico
TicAc	<i>Ticks Acima</i>
TicAB	<i>Ticks Abaixo</i>
USD	Dólar

Sumário

1	Introdução	14
2	Séries Temporais	17
2.1	As Séries Temporais	17
2.2	Estacionariedade	18
3	O Mercado FOREX, Inteligência artificial e Predição	21
3.1	O Mercado FOREX, Pares de Moedas e suas Operações	21
3.2	Predição	23
3.3	Análise de Séries Temporais	24
3.4	Inteligência Artificial na Previsão de Séries Financeiras	25
3.5	Redes Neurais Artificiais	25
3.6	A Arquitetura da Rede Neural Recorrente	28
4	Modelo Proposto de Predição	31
4.1	Metodologia	33
4.2	Coleta de dados e pré-processamento	34
4.2.1	A Análise Técnica	34
5	A Rede Neural	39
5.1	LSTM	39
5.2	Definição dos Hiper-parâmetros	39
5.3	Preparação dos Dados para a rede LSTM	40
5.4	Estratégia de Treinamento	41
5.5	Estratégia de Teste	42
5.6	Interpretação das saídas da rede	43
6	Análise dos Resultados	44
7	Conclusões	48
7.1	Discussão e Propostas	48
	Referências	50

Apêndices	53
APÊNDICE A Ambiente de desenvolvimento da rede neural utilizada	54

1 Introdução

O ser humano tem sido curioso sobre o futuro desde que tomou consciência deste. Por muito tempo as pessoas têm tentado prevê-lo das mais variadas maneiras, obtendo resultados diversos. Nesse contexto, a análise e previsão de séries temporais financeiras é uma área de estudo popular, importante e atualmente em constante desenvolvimento dentro da ciência da computação. A ideia de que computadores podem auxiliar na previsão de ativos do mercado financeiro desperta interesse de pesquisadores e analistas, que enxergam nesta ideia a possibilidade de solução de desafios científicos além de ganhos financeiros.

O mercado financeiro pode ser analisado graças ao conjunto de dados que é gerado, e posteriormente coletado, a partir da oscilação de uma ou mais variáveis. Este conjunto é classificado como uma série temporal, que por definição, é uma sequência de observações obtidas em intervalos regulares de tempo durante um período específico [Yao e Tan \(2000\)](#). Nesta perspectiva, a análise de séries temporais através, do crescente poder de processamento da computação moderna, tem se configurado como uma área bastante promissora e com grande impacto técnico/científico [Dhamija e Balla \(2010\)](#). Esta, ainda é considerada um domínio de grande importância visto que o conhecimento obtido através de sua aplicação, tem potencial para auxiliar na construção de ferramentas estratégicas a serem utilizadas tanto em nível macro quanto microeconômico, tendo como principal finalidade o desenvolvimento de modelos que representem o comportamento dos valores da série, tendo em vista, entre outros objetivos, a sua previsão.

A previsão de uma série temporal só é possível se houver um sistema causal relativamente constante, relacionado com o tempo, que tenha exercido influência sobre os dados no passado e que mantenha sua ação sobre os dados no futuro [Shumway e Stoffer \(2010\)](#). Dito isso, a previsão de valores numa série temporal originada do mercado financeiro se torna potencialmente viável já que atende a determinada condição, possibilitando a elaboração de planos a longo, médio e curto prazo, além de mostrar-se como opção auxiliar aos investidores no processo de tomada de decisões, que consequentemente ampliam a possibilidade de ganhos no maior e mais líquido dos mercados financeiros, o Foreign Exchange também conhecido como FOREX, com trilhões de dólares negociados diariamente [Bortoli \(2015\)](#).

O FOREX é composto por duas séries temporais que sinalizam os valores das cotações para compra e venda de um determinado par de moedas. Estas séries possuem características estocásticas e não lineares, tornando a tarefa de previsão mais complexa e desafiadora. A complexidade do problema atrelada a possibilidade de ganhos financeiros

reais tem fomentado o estudo e desenvolvimento de modelos de previsão que se aproximem de resultados cada vez mais precisos. Um dos mais conhecidos e utilizados é o modelo auto-regressivo integrado de média móvel (*ARIMA*) [Box, Jenkins e Reinsel \(1994\)](#). Este modelo parte do pressuposto de que a série temporal é linear e segue uma distribuição estatística específica. Entretanto, limitações importantes se tornam explícitas quando se sabe que o comportamento não linear das séries originadas de problemas práticos não se encaixa adequadamente neste modelo, necessitando assim de uma abordagem que permita a realização de mudanças de acordo com a estrutura de mercado.

Nos últimos anos, a Inteligência Artificial (IA) vem auxiliando a construção de modelos e sistemas capazes de analisar grandes quantidades de dados em alta velocidade e melhorar a si mesmos através de tal análise. Nesse contexto, algumas abordagens surgiram como forma de alcançar objetivos específicos através da combinação da computação e da IA. A fim de solucionar certas limitações dos modelos lineares, como por exemplo a pressuposição de que existe uma relação linear entre as variáveis de uma série temporal, as redes neurais artificiais (RNAs) passaram a atrair atenção para área de previsão desse tipo de série. Diferente do modelo *ARIMA*, a RNA não parte de pressupostos estatísticos para a modelagem. O processo é feito de forma adaptativa e não linear, baseando-se nos dados de entrada¹ recebidos pela rede. A capacidade de aprendizagem por meio de exemplos e de generalização da informação aprendida é uma das características mais atraentes das redes neurais.

Trabalhos na área indicam que até mesmo redes neurais simples podem ter uma eficiência similar e em muitos casos significativamente melhor na previsão de séries temporais se comparadas a outros modelos [Adhikari e Agrawal \(2013\)](#). O trabalho de [Dhamija e Balla \(2010\)](#) comparou a acurácia em relação a capacidade de previsão das redes neurais e dos modelos ARCH, chegando a conclusão que ambos podem ser usados para fazer previsões de séries do mercado financeiro. [Sharda e Patil \(1992\)](#) compararam redes neurais com o modelo ARIMA utilizando dados de séries temporais numa competição conhecida como *M-competition*. Os resultados mostraram que um modelo de rede neural simples testado no subconjunto proveniente das séries previu tão bem quanto o sistema Box-Jenkins que aplica o modelo *ARIMA*. [Yao e Tan \(2000\)](#) chegaram a conclusão que até mesmo sem um conhecimento amplo e sem o uso extensivo de dados reais do mercado FOREX é possível utilizar-se das redes neurais para fazer previsões úteis e obter ganhos teóricos significativos. [Kondratenko e Kuperin \(2003\)](#) provaram de forma empírica que o uso de indicadores categorizados como rastreadores de tendência podem ser utilizados como dados de entrada numa rede neural recorrente, tornando esse modelo aplicável para previsão de séries temporais do mercado FOREX.

¹ Esses dados de entrada podem ser entendidos como a percepção que a rede neural tem em relação as várias formas de informação obtidas do ambiente exterior.

Embora diversos trabalhos tenham abordado a utilização de RNAs para melhorar a previsão de valores em séries do mercado financeiro, as abordagens vistas frequentemente pecam por não incluir em suas análises as variáveis² que modelam o comportamento real das operações financeiras sob as séries. Além disso, há carência em relação a trabalhos que abordem o desenvolvimento de novos indicadores utilizados em conjunto com arquiteturas de redes neurais, a fim de obter-se melhores taxas efetivas de acerto³ e possíveis lucros reais. Neste sentido, trabalhos prévios desenvolvidos pelo grupo de pesquisa da Universidade do Estado da Bahia iniciaram um novo esforço na tentativa de incluir essas variáveis na análise de predição [Bortoli \(2015\)](#), chegando a taxas efetivas de acerto na ordem de 42%.

Sabe-se que um dos principais objetivos das aplicações financeiras é a obtenção de lucros. Com a intenção de alcançar este propósito, o aperfeiçoamento constante de modelos que possam auxiliar na previsão de valores de séries temporais do mercado financeiro segue sendo a principal contribuição desse trabalho, visto que uma ferramenta mais eficaz na predição desse tipo de série poderia ser aplicada também em outras áreas, onde a tomada de decisões se baseie no comportamento de uma série temporal.

Assim sendo, o objetivo desse trabalho é propor um novo indicador, baseando-se na análise técnica dos dados do mercado FOREX, comparando-o com o indicador de [Bortoli \(2015\)](#), visando sua utilização para melhoria na taxa de previsibilidade de ativos neste mercado através do uso de técnicas da inteligência artificial. O presente trabalho está dividido em 5 capítulos.

Os capítulos 2 e 3 constituem o referencial teórico deste trabalho. O capítulo 2 aborda o conceito de série temporal e estacionariedade. O capítulo 3 descreve o funcionamento do mercado FOREX, a análise de suas séries temporais, o emprego da inteligência artificial na previsão de séries financeiras e as redes neurais artificiais.

O capítulo 4 traz o modelo proposto de predição, a metodologia empregada, a coleta dos dados e a análise técnica aplicada a estes dados do mercado FOREX.

A arquitetura de rede neural utilizada, assim como os seus parâmetros, estratégias de treinamento e teste são abordados no capítulo 5.

Os resultados são coletados e comparados no capítulo 6, onde o desempenho da rede é avaliado em conjunto com o indicador criado

Finalmente, o capítulo 7 traz as conclusões e os trabalhos futuros.

² Comissões, permanência, spreads, diferenças cambiais

³ Entende-se por taxas efetivas de acerto as previsões que geram resultado financeiro, ou seja, não só as previsões de tendências assertivas mas a inclusão nestas previsões dos valores que devem ser descontados e que fazem o saldo após a transação tornar-se positivo ou negativo.

2 Séries Temporais

Este capítulo traz o conceito, estrutura e características das séries temporais além da definição de estacionariedade e seus processos.

2.1 As Séries Temporais

[Adhikari e Agrawal \(2013\)](#) explicam que uma série temporal é uma sequência de observações obtidas em intervalos regulares de tempo durante um período específico. Estas séries podem ser representadas matematicamente como sendo um conjunto de vetores x_1, x_2, \dots, x_T ou $x_t, t = 1, 2, \dots, T$ onde T representa o tempo decorrido [Cochrane \(2005\)](#). Os autores ainda explicam que as séries temporais podem apresenta-se numa diversidade de áreas como a engenharia, meteorologia, medicina, e que, a depender de sua origem, podem ser modeladas a partir de funções matemáticas. Contudo, existe um tipo de série temporal frequentemente originária do mercado financeiro e que não possui função definida capaz de descrever seu comportamento ao longo do tempo, as séries temporais estocásticas. Além de se distinguirem em relação a sua origem, as séries temporais ainda possuem características que as classificam da seguinte forma segundo [Adhikari e Agrawal \(2013\)](#):

1. **Univariadas ou Multivariadas:** Quando uma série contém registros de uma única variável é denominada como univariada, como por exemplo, os valores diários do preço das ações em uma bolsa de valores. Se os registros de mais de uma variável são considerados, então sua denominação torna-se multivariada, a exemplo das vendas e gastos com propagando semanais de um produto.
2. **Contínuas ou Discretas:** Uma série é contínua quando a observação de seus valores é feita de forma contínua no tempo, como por exemplo, o registro de um eletrocardiograma de uma pessoa. Ao passo que, uma série é classificada como discreta quando contém observações medidas em pontos de tempo específicos, geralmente equiespaçados, como o número mensal de casos notificados de febre amarela.

Uma série temporal possui componentes que são geralmente utilizados para fazer sua análise. Esta análise é descrita como a procura de autocorrelação num conjunto de dados extraídos de uma determinada série, tendo a finalidade de compreender tanto sua estrutura básica quanto a função que a representa de forma aproximada, sendo tradicionalmente baseada na decomposição da série temporal em pelo menos três componentes. Tendência, ciclo e sazonalidade [Adhikari e Agrawal \(2013\)](#). O componente tendência indica o comportamento da série em relação à subida, descida ou estagnação durante um

período de tempo. A sazonalidade indica a repetição de um determinado padrão dentro de uma série no período de um ano. O ciclo exerce uma indicação análoga a sazonalidade, com a diferença de que o período analisado é superior a um ano.

Além disso, sabe-se que as séries temporais possuem em sua essência a característica estocástica e que sua previsão é baseada na probabilidade, portanto um modelo matemático que descreva sua estrutura probabilística baseando-se em valores passados é descrito como sendo o processo estocástico [Hipel e McLeod \(1994\)](#). Os autores ainda explicam que empregar um modelo apropriado numa série temporal visando o uso em aplicações é chamado mais especificamente de análise de séries temporais. Esta análise tem como principal finalidade o desenvolvimento ou a utilização de modelos que representem o comportamento dos valores de uma série temporal, tendo em vista, entre outros objetivos, a sua predição.

Pode-se observar na figura 1, o fragmento de uma série temporal extraída do mercado financeiro FOREX, onde apresenta-se os valores de cotação para o par de moeda Euro/Dólar representado pelas siglas EUR/USD, significando as cotações onde, 1 euro vale em cada instante de tempo uma quantidade de dólar. A exemplo, quando no dia 22 de dezembro as 14:00 horas o valor de 1.00 euro comprava aproximadamente 1.04390 dólares.

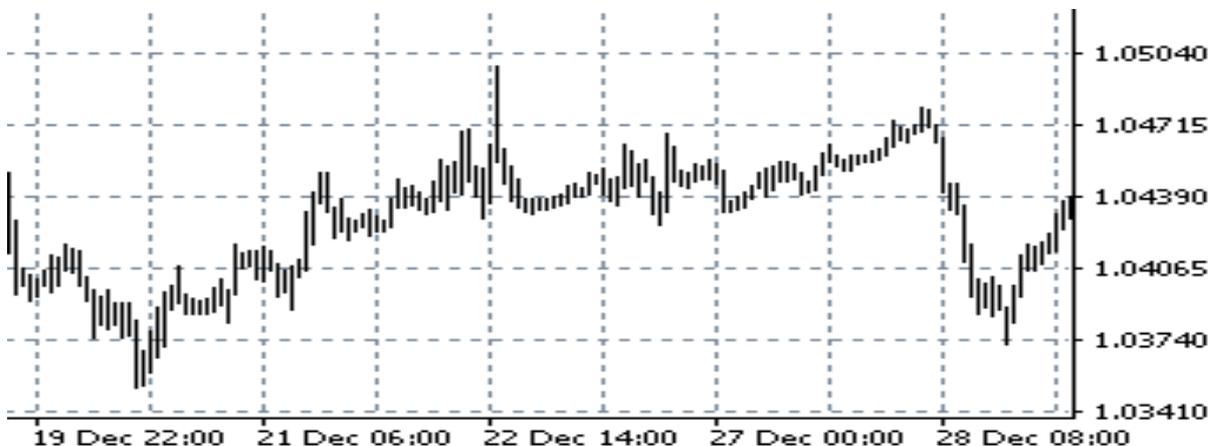


Figura 1: Série temporal do mercado FOREX retirada da ferramenta MetaTrader - EXNESS referente ao par de moeda EUR/USD no período de 19 a 28 de dezembro de 2016

As séries temporais ainda possuem um certo equilíbrio estatístico que é explicado pela estacionariedade, cuja existência possibilitou que grande parte dos recursos de predição atuais fossem elaborados.

2.2 Estacionariedade

[Adhikari e Agrawal \(2013\)](#) dizem que uma série temporal é considerada estacionária quando há um desenvolvimento ao longo do tempo de forma aleatória em torno

de uma média constante apresentando alguma forma de equilíbrio estável. Para que seja possível a construção de modelos úteis para a previsão de séries temporais é necessário que propriedades estatísticas tais como média e variância de um processo estacionário não dependam do tempo [Hipel e McLeod \(1994\)](#). Então, para a simplificação da construção dos modelos supracitados e redução da complexidade matemática destes é que se espera que as séries sejam estacionárias, quando pretende-se utilizar seus dados para realizar previsões. Porém como descrito por [Adhikari e Agrawal \(2013, tradução nossa\)](#) "Quanto maior for o tempo de observações históricas, maior será a chance de que as séries temporais exibam características não estacionárias"¹. Nestes casos, é comumente empregada a técnica de normalização nos dados extraídos das séries temporais através do uso de famílias de funções matemáticas, visando a eliminação de possíveis tendências ou padrões sazonais, tornando o fragmento de série observado estacionário.

Ainda dentro desse conceito, existem dois tipos de processos estacionários que precisam ser definidos. O primeiro deles é o fortemente estacionário. Neste, todo conjunto de variáveis aleatórias do processo permanece constante não importa qual seja o ponto de medição. O segundo deles ocorre com mais frequência e é chamado de fracamente estacionário. Segundo [Gujarati e Porter \(2009\)](#), este só é verdade se e somente se a média for constante, a variância for constante e a covariância for dependente apenas do intervalo entre dois períodos de tempo.

Pode-se ver mais a esquerda da figura 2, o exemplo de uma série temporal genérica não estacionária com tendência crescente. Esta mesma série pode ser vista em sua forma estacionária mais a direita. Para que seja possível alcançar o estado estacionário, transformações através da utilização de técnicas matemáticas, como por exemplo, a aplicação da diferenciação, se faz necessária. Veja que quando estacionária, grande parte das observações feitas ao longo do tempo variam em torno de um intervalo bem definido.

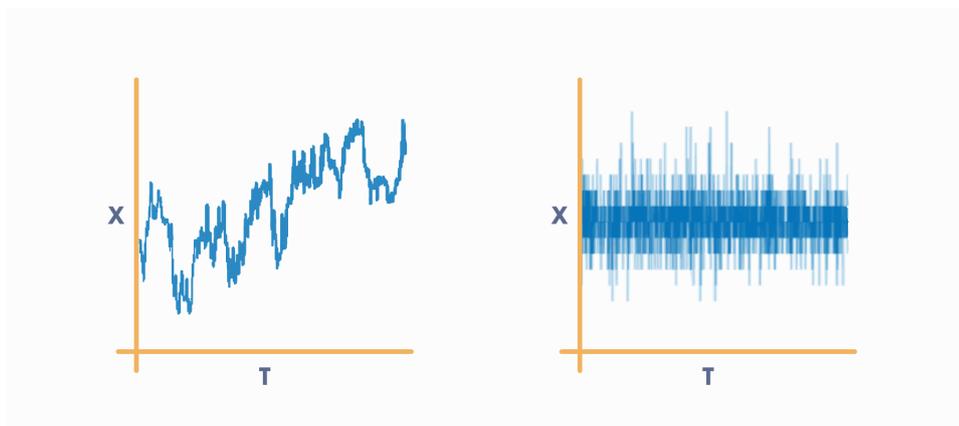


Figura 2: Série temporal genérica

¹ do texto original: The greater the time span of historical observations, the greater is the chance that the time series will exhibit non-stationary characteristics

O uso da estacionariedade fraca é comum na grande maioria das vezes, ao ponto que quando houver referência a um processo estacionário sem uma definição mais específica, pode-se assumir que este é fracamente estacionário.

Neste trabalho, as séries temporais utilizadas para análise e obtenção de dados são sempre originárias do FOREX. Este apresenta como uma de suas principais características um movimento desordenado e volátil, produzindo séries temporais com as propriedades já mencionadas e que podem ser utilizadas para prever o comportamento futuro deste mercado.

3 O Mercado FOREX, Inteligência artificial e Predição

3.1 O Mercado FOREX, Pares de Moedas e suas Operações

O mercado de câmbio também conhecido como Forex, é o maior e o mais líquido dos mercados financeiros, com trilhões de dólares negociados diariamente [Bortoli \(2015\)](#). É considerado um mercado global descentralizado, sendo utilizado para realizar transações de compra e venda de moedas estrangeiras. O FOREX não é um único mercado em si, mas sim uma rede global de mercados que conectam investidores de todo o mundo [Aloud et al. \(2012\)](#). Esses investidores podem ser governos, bancos centrais ou até mesmo investidores individuais. Devido aos avanços nas tecnologias de comunicação, principalmente as relacionadas com aplicações *web*, mudanças significativas ocorreram neste, que passou a receber um número grande de investidores e suas respectivas plataformas de comércio baseadas predominantemente na internet, também conhecidas como (*FOREX market maker*) ou formadores de mercado FOREX.

[Aloud et al. \(2012\)](#) definem essas plataformas formadoras de mercado como instituições que fornecem liquidez para um número de pares de moedas e estimam um preço, tanto para compra quanto para venda dentro de uma plataforma, numa tentativa de obter lucros através da diferença entre o *bid* e o *ask* chamada de *spread*. Conforme [Bortoli \(2015\)](#), "O *ask* é o valor que o mercado está disposto a vender uma moeda e sempre será um valor maior ou igual ao *bid*. Já o *bid* é o valor que o mercado está disposto a comprar uma moeda". Esse valor é ajustado pelas plataformas de acordo com a procura e oferta do mercado, notícias e preços de outras plataformas, entre outros fatores. O mercado FOREX é bastante peculiar em seu funcionamento se o compararmos ao mercado de ações. Ele funciona 24 horas por dia, dos domingos às sextas-feiras. Apesar de funcionar ininterruptamente durante 5 dias da semana, as operações realizadas pelos *traders* não seguem um comportamento uniforme, tendo picos no momento em que os maiores operadores desse mercado financeiro iniciam suas transações.

Tais transações baseiam-se em pares de moedas, onde cada moeda é representada pela combinação de três letras como por exemplo EUR/USD (Euro/Dólar). Este par representa respectivamente uma base e uma cotação (base/cotação) onde ambas as moedas são negociadas. Para facilitar o entendimento, pode-se pensar numa operação do mercado FOREX como sendo uma aposta de um investidor em uma das moedas em relação ao seu valor de mercado num determinado momento futuro. Consideremos o par EUR/USD do exemplo anterior. Em um instante específico, com 1.00 EUR pode-se comprar 1.3940

USD. Essa taxa de câmbio pode mudar a qualquer momento por fatores já mencionados, causando as flutuações de câmbio. Considerando os fatores externos, pode-se acreditar que o EUR irá se tornar ainda mais forte que o USD, o que significa dizer que com 1.00 EUR poderá se comprar mais unidades de USD. Suponha-se a aposta de 10.000 EUR que, no instante inicial e utilizando a taxa de cambio mencionada acima, pode comprar 13.940 USD. Considerando-se também que a aposta em questão levou 5 minutos e que durante esse tempo a taxa cambial do par EUR/USD passou de 1.3940 para 1.3947, no instante final tem-se que 10.000 EUR equivalem a 13.947 USD. Em apenas 5 minutos foi obtido o lucro de 7 USD. Essa diferença ao final da transação é geralmente expressa pela menor unidade usada para representar a mudança de preço numa taxa de câmbio, chamada de *pip* (*Percentage in point ou price interest point*). Nesse caso, a transação obteve como resultado o ganho de 7 *pips* Zulu (2010).

Um outro conceito bastante utilizado quando deseja-se apostar no mercado FOREX e não dispõe-se de grandes recursos para tal, é o de alavancagem (*leverage*, em inglês). Se o mesmo investidor que acreditava que o EUR se tornaria mais forte que o USD passar a acreditar que esse crescimento será muito mais forte ao ponto de valer o investimento de um montante consideravelmente maior, como por exemplo 100.000 EUR, é possível fazer essa aposta, ainda que não se possa banca-la completamente. Isso só é possível pelo fato de as plataformas, também conhecidas como corretoras, possibilitarem o empréstimo necessário para que se assuma o risco. É necessário apenas o que é chamado de depósito colateral (*collateral deposit*) que na prática serve para cobrir quaisquer que sejam os riscos de crédito possíveis no decorrer da execução de operações cambiais Zulu (2010). É possível fazer empréstimos de até centenas de vezes o valor do depósito. Por exemplo, se um investidor com um depósito de 1.000 EUR fizer uma aposta do mesmo valor, significa dizer que não há alavancagem ou a razão pode ser representada por 1:1. Caso decida-se fazer com esse mesmo depósito uma aposta de 50.000 EUR, então a razão da alavancagem sobe para 50:1.

Uma das estratégias utilizadas para negociação no mercado FOREX se dá através do uso de *Expert Advisors*, que são programas desenvolvidos especificamente para automatizar os processos analíticos e de negociação. Estes programas seguem estratégias previamente implementadas, possibilitando por exemplo que um investidor sem um conhecimento profundo da análise técnica ou fundamental possa negociar, já que todos os cálculos necessários estão incluídos no programa. Entre as estratégias automatizadas mais comuns está a possibilidade da definição de paradas em momentos oportunos, como por exemplo o fechamento de uma ordem caso atinja-se um nível de prejuízo pré-definido, chamado de *Stop-Loss* e realização da mesma ação caso ocorra o inverso, atinja-se um nível de ganho, chamado de *Take-Profit*.

3.2 Predição

A predição pode ser definida como o ato ou efeito de predizer, de afirmar o que vai acontecer no futuro. Fora do contexto científico, o termo predição é frequentemente usado para se referir a uma opinião ou opinião informada, caso a pessoa responsável seja experiente no campo que se deseja prever. No mercado financeiro, predizer significa tentar determinar o valor futuro de um ativo negociável¹.

É razoável e natural classificar a qualidade de uma previsão pela sua acurácia. Sabe-se que a quantidade de informação relacionada com o passado do que se deseja prever combinada com o intervalo de predição que se tem a intenção de observar num ponto futuro é crucial para obter-se resultados satisfatórios, independente do modelo de previsão empregado no processo [Diebold e Kilian \(2001\)](#). As séries temporais econômicas utilizadas como base para previsão de ativos no mercado financeiro, ainda contam com algumas medidas utilizadas para estimar seus níveis de previsibilidade. [Diebold e Kilian \(2001\)](#), propuseram uma medida baseada na razão entre a perda esperada numa previsão de curto prazo e a perda esperada numa previsão de longo prazo. Essa medida de previsibilidade é importante, pois além de outros aspectos, permite o casamento mais adequado entre o conjunto de dados que pretende-se usar e do modelo de previsão a ser empregado. Para um melhor entendimento, se a série temporal A tem maior previsibilidade que uma série temporal B, as previsões feitas em A tendem a ter erros menores quando comparadas as previsões feitas em B. Trazendo para o contexto de aplicações financeiras, conhecer a previsibilidade minimiza os riscos e maximiza o retorno das decisões de investimento.

O objetivo principal da predição quando se trata de séries temporais do mercado FOREX não é saber exatamente o valor de um ponto num futuro definido, sobretudo porque o comportamento destas séries é estocástico, mas sim basear-se na probabilidade partindo do pressuposto de que a tendência de variações continuará a se repetir no futuro, sendo possível a sua previsão de forma aproximada tanto quanto possível. O problema da predição de séries temporais é encontrado em muitas áreas de atuação e tem como ponto comum o questionamento de quão viável é analisar e usar dados obtidos num determinado período de tempo passado para prever o futuro. De acordo com [Kuna \(2015\)](#), o horizonte de previsão pode ser definido como sendo o número de intervalos de tempo a frente do ponto atual para o qual pretende-se fazer a previsão. Além disso, sabe-se que um dos maiores desafios no campo da predição é o emprego de técnicas e modelos que se adequem de forma mais satisfatória quando é necessário utilizar-se de um horizonte de previsão com um período de tempo mais longo, haja visto que a qualidade dos modelos existentes para previsão são frequentemente mais eficientes para horizontes com um número menor de pontos observados. Abaixo são apresentados os dois métodos mais vistos atualmente para

¹ Esses ativos podem ser dinheiro, um direito contratual para entregar ou receber dinheiro ou outro tipo de instrumento financeiro

prever o comportamento do mercado FOREX.

3.3 Análise de Séries Temporais

Embora bastante diferentes, os dois tipos de análise mais utilizados nas séries temporais, a análise técnica e a análise fundamentalista, têm como principal objetivo auxiliar na previsão de séries temporais. De uma forma geral os adeptos da análise técnica estudam os efeitos do movimento do mercado enquanto que os fundamentalistas dão maior ênfase a causa. A análise técnica é definida por [Murphy \(1999\)](#) como sendo o estudo da ação do mercado, principalmente através do uso de gráficos, com o propósito de prever tendências nos preços futuros de ativos. Esta análise ainda tem três princípios essenciais segundo o autor:

1. **O mercado desconta tudo:** Isto significa que o preço real é um reflexo de tudo que pode afetar o mercado de alguma maneira, como por exemplo, a oferta e procura, fatores políticos e etc.
2. **Preços movem-se em tendências:** O conceito de tendência é essencial quando se fala em análise técnica. Esta é utilizada para identificar padrões de comportamento de mercado, com o objetivo de fazer operações baseando-se na direção da tendência encontrada.
3. **A história se repete:** Os padrões gráficos do mercado Forex vêm sendo reconhecidos e classificados para mais de um século e a maneira pela qual muitos padrões se repetem, formando inclusive figuras, levam à conclusão de que a psicologia humana muda pouco e que se esses padrões funcionaram no passado eles vão continuar funcionando no futuro.

A análise fundamentalista por sua vez foca na força econômica da oferta e procura, responsável por fazer os preços subirem, caírem ou permanecerem no mesmo patamar [Murphy \(1999\)](#). Esta abordagem examina todos os fatores tidos como relevantes e que afetem o valor do preço do mercado com a intenção de determinar seu valor intrínseco. Ambas as abordagens de previsão para o mercado tentam resolver o mesmo problema que é o de determinar a direção provável que o seu preço tende a seguir, elas apenas o encaram de formas diferentes.

Tendo em vista a direção provável que o preço tende a seguir, a quantidade de valores a serem considerados e as diversas variáveis que têm influência sobre o mercado, uma das abordagens encontradas na literatura para solucionar esse problema seria o uso das técnicas de inteligência artificial.

3.4 Inteligência Artificial na Previsão de Séries Financeiras

A quantidade de dados produzidos e armazenados de forma digital vem crescendo de forma constante, chegando a casa dos bilhões de gigabytes diários [Wall \(2014\)](#). Nos últimos anos, a Inteligência Artificial (IA) vem auxiliando a construção de modelos e sistemas capazes de analisar esse grande volume de dados em alta velocidade e melhorar a si mesmos através de tal análise sendo aplicados com sucesso numa diversidade de áreas como a medicina, geologia, física e engenharia. A possibilidade de ganhos reais através de uma ferramenta útil que seja capaz de prever com certa acurácia mudanças no preço de ativos financeiros é um grande motivador para o número cada vez maior de pesquisadores que se interessam pela área de previsão de séries temporais do mercado FOREX. Dentre as técnicas de IA mais utilizadas para modelagem de sistemas para esse mercado, pode-se citar a que é provavelmente uma das mais exploradas, as Redes Neurais Artificiais (RNA) [Dhamija e Balla \(2010\)](#), [Haykin \(1998\)](#), [Bortoli \(2015\)](#).

3.5 Redes Neurais Artificiais

No ramo da previsão de séries temporais, alguns modelos ganharam popularidade. Um dos mais conhecidos e utilizados é o modelo auto-regressivo integrado de média móvel (*ARIMA*) [Box, Jenkins e Reinsel \(1994\)](#). Este modelo parte do pressuposto de que a série temporal é linear e segue uma distribuição estatística específica. Sua popularidade se dá pela versatilidade em representar várias séries temporais de forma simplificada. Entretanto, limitações importantes se tornam explícitas quando se sabe que o comportamento não linear das séries originadas de problemas práticos não se encaixa adequadamente neste modelo, necessitando assim de uma abordagem que permita a realização de mudanças de acordo com a estrutura de mercado.

A fim de solucionar certas limitações dos modelos lineares, as redes neurais artificiais (RNA's) têm atraído atenção para a área de previsão de séries temporais. Estas redes foram concebidas para imitar as características biológicas do neurônios no cérebro humano e do sistema nervoso [Haykin \(1998\)](#). Diferente do modelo ARIMA, a RNA não parte de pressupostos estatísticos para a modelagem. O processo é feito de forma adaptativa e não linear, baseando-se nos dados de entrada recebidos pela rede. A capacidade de aprendizagem por meio de exemplos e de generalização da informação aprendida é uma das características mais atraentes das redes neurais. Para um melhor entendimento, uma rede neural artificial opera criando conexões com um conjunto de diferentes elementos de processamento, sendo análogo aos neurônios em um cérebro biológico, recebendo e retransmitindo informações. O primeiro modelo de neurônio artificial foi criado por [McCulloch e Pitts \(1943\)](#) e está representado de forma adaptada por ([Bortoli \(2015\)](#)) na figura 3.

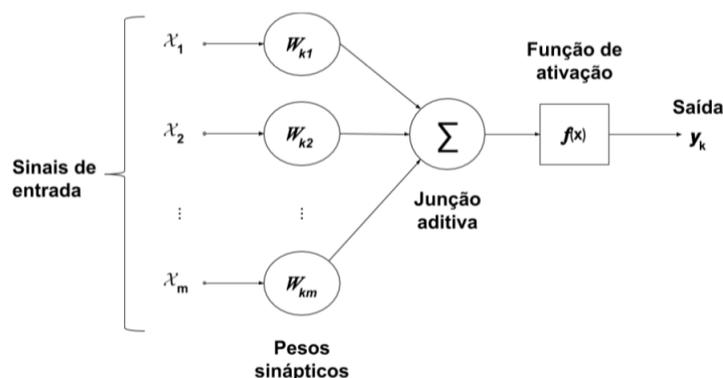


Figura 3: Modelo de Neurônio Artificial (Adaptado de: Haykin (2001, p. 36))

As RNA's são caracterizadas por componentes que podem ser classificados como físicos e não-físicos: Os componentes físicos são os elementos de processamento e as conexões. Se fizermos uma analogia com um grafo orientado (dígrafo)², pode-se dizer que os nodos são elementos de processamento e as arestas as conexões. Já os componentes não-físicos são chamados de padrões e funções. De acordo com Simpson (1990) os padrões são os dados de entrada da rede ao passo que as funções são os modelos matemáticos utilizados no treinamento e reconhecimento de padrões.

Para alcançar um bom desempenho em qualquer que seja a tarefa, as redes neurais empregam uma interconexão maciça entre os relativamente simples componentes computacionais, denominados neurônios ou unidades de processamento. Portanto, Haykin (1998) apresenta a seguinte definição para uma rede neural quando vista como uma máquina adaptativa:

“Uma rede neural é um processador massivamente paralelo e distribuído, composto de unidades de processamento simples, que tem uma propensão natural para armazenar conhecimento adquirido pela experiência, tornando-o disponível para uso.”

Uma RN ainda assemelha-se ao cérebro em dois aspectos:

1. O conhecimento adquirido pela rede é originário do seu ambiente através de um processo de aprendizagem
2. A força de conexão entre os neurônios, conhecida como pesos sinápticos, são usados para armazenar o conhecimento adquirido

O procedimento utilizado para executar o processo de aprendizagem é feito através dos chamados algoritmos de aprendizagem. Um dos mais bem sucedidos no que diz a respeito ao treinamento de RNs é o *Stochastic Gradient Descent* ou *SGD*. Frequentemente

² Um dígrafo tem arestas entre nodos com um só sentido. Isso faz com que a informação siga numa direção definida

utilizado para o treinamento de redes neurais com múltiplas camadas, o *SGD* não demanda a utilização de todo o conjunto de treinamento de uma única vez pelo modelo a ser treinado, visto que do ponto de vista computacional, utilizar apenas uma etapa para percorrer todo o conjunto de treinamento, para então poder atualizar os pesos, pode ser custoso, em especial, se o conjunto for relativamente grande. Com o *SGD*, é possível fragmentar o conjunto de treinamento, possibilitando assim a atualização dos pesos da rede a cada iteração, ou seja, cada fragmento do conjunto de treinamento utilizado no processo de aprendizagem faz com que os pesos sejam ajustados. Técnicas computacionais como a *backpropagation*, também são utilizadas no processo de aprendizagem. Esta tem como princípio a modelagem de uma determinada função, a partir da modificação dos pesos internos dos sinais de entrada visando produzir um sinal de saída esperado, tornando a atualização dos pesos sinápticos mais eficiente. Os algoritmos mencionados são, em muitos casos, utilizados em conjunto para treinar uma rede neural.

No caso deste trabalho e já considerando as propriedades do mercado FOREX, é importante ainda a definição da função de ativação que seja mais apropriada, tanto em relação ao problema abordado quanto a arquitetura de rede neural utilizada. As funções de ativação são importantes, pois têm o objetivo de incorporar a não linearidade a rede, permitindo que estas possam produzir comportamento complexo. Dentre as funções mais utilizadas em redes neurais com múltiplas camadas, estão a tangente hiperbólica 3.1 e a sigmóide 3.2.

$$\tanh(t) = \frac{e^t - e^{-t}}{e^t + e^{-t}} \quad (3.1)$$

$$S(t) = \frac{1}{1 + e^{-t}} \quad (3.2)$$

Através da utilização destas funções nas rede neurais, é possível obter-se de forma aproximada uma função que corresponda de maneira mais acurada a curva existente no conjunto de dados utilizado. Essa aproximação acontece principalmente através da regressão não linear em conjunto com a aplicação das funções previamente mencionadas Zeviani, Ribeiro Jr e Bonat (2013).

Dentre os tipos de redes mais aplicados para problemas de regressão não linear e predição de séries temporais, estão as Redes Neurais Recorrentes (RNRs). Apesar de não serem as únicas com a capacidade de representação de tempo, são as mais utilizadas e testadas para problemas desse tipo Kondratenko e Kuperin (2003), Kuna (2015), Persio e Honchar (2016).

3.6 A Arquitetura da Rede Neural Recorrente

Olah (2015) descreve uma RNR como um tipo de rede não tradicional. Sua capacidade particular de manter a informação prévia se encaixa melhor em problemas relacionados com sequências e listas, problemas estes que não podem ser facilmente solucionados com redes neurais tradicionais. As redes neurais recorrentes ainda contam com um loop em sua estrutura, principal responsável pela sua diferenciação.

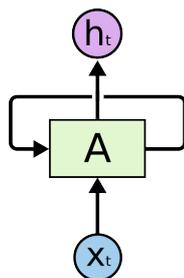


Figura 4: Arquitetura básica de uma RNN e seu loop (Olah (2015))

A figura 5 exemplifica o que é parte de uma RNR chamada A, recebendo como entrada X_t e retornando o valor de saída H_t . Este loop é o que transmite a informação de um ponto da rede para outro. Apesar de aparentemente mais complexa, uma RNR pode ser melhor explicada quando pensa-se de forma convencional, levando em consideração que uma RNR nada mais é que múltiplas cópias de uma mesma rede, cada uma destas passando informação para sua sucessora Olah (2015), como pode-se observar abaixo:

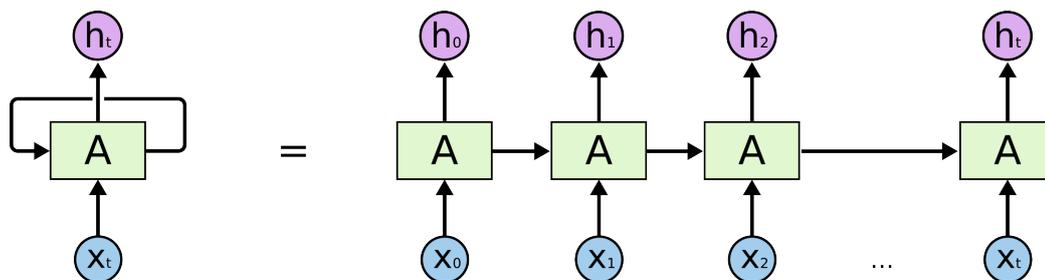


Figura 5: Representação da arquitetura RNN com seu loop expandido (Olah (2015))

É difícil citar RNRs sem mencionar uma das arquiteturas mais utilizadas e que atualmente tem dado resultados significativamente melhores quando comparados ao modelo padrão de rede neural recorrente, as redes *Long Short Term Memory*. Também conhecida como *LSTM*, essa arquitetura particular de RNR foi apresentada por Hochreiter e Schmidhuber (1997) visando resolver principalmente o problema das dependências de longo prazo.

Tomemos o exemplo a seguir para um melhor entendimento do problema em questão e o que as *LSTM's* se propõem a resolver. Considere um modelo linguístico que tem como objetivo prever palavras baseando-se em termos anteriores numa frase. Ao tentar-se

prever a última palavra em "Os peixes vivem na *água*", sem um contexto muito mais elaborado e complexo é bastante fácil imaginar que a palavra a ser prevista é água. Em casos onde a distância entre a informação relevante e o espaço que precisa ser preenchido é pequeno, as RNRs podem aprender com o uso da informação relevante próxima. Porém, existem casos onde há necessidade de se analisar um contexto mais complexo para realizar o mesmo tipo de predição. Considere a frase "Eu nasci em Portugal... eu falo fluentemente *português*". A informação contextual sugere que a palavra a ser predita é o nome da língua, mas se quisermos especificar que língua seria essa é preciso recorrer ao contexto da palavra Portugal, que está notavelmente mais distante do que o caso do primeiro exemplo. Essa distância entre a informação relevante e o ponto de predição pode ser muito maior que esses casos, dificultando todo o processo. Com o crescimento dessa distância, as RNRs tradicionais passam a não conseguir aprender através da conexão da informação, o que não acontece com as redes *LSTM's*, já que estas foram pensadas primordialmente para evitar este problema caracterizado como dependências de longo prazo Olah (2015).

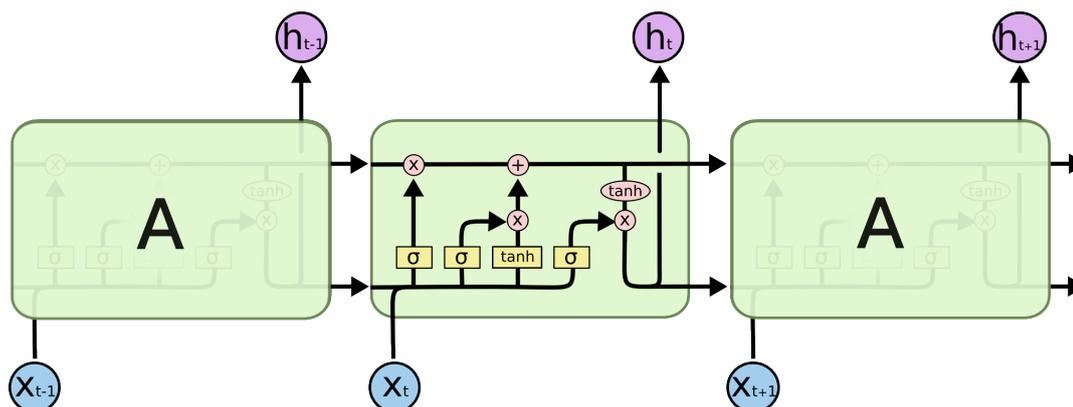


Figura 6: Representação interna de um módulo que compõe a arquitetura LSTM Olah (2015))

De forma simplificada, pode-se ver na figura acima um bloco *LSTM*. Esse tipo de rede tem a capacidade de remover ou adicionar informação numa estrutura chamada *cell state*, representada pela linha horizontal que passa por toda parte superior do diagrama. Para questão de entendimento, pode-se fazer analogia a *cell state* com uma correia de transporte, onde a informação trafega passando apenas por algumas interações lineares como operações de soma e multiplicação. Essa capacidade da rede é regulada por outra estrutura, os *gates* ou portões. *Gates* oferecem uma forma opcional de tráfego da informação. Eles são compostos de uma camada de rede neural sigmóide, representada com o retângulo amarelo com o símbolo sigma, e uma operação de multiplicação. Um bloco LSTM tem três desses *gates*, para proteger e controlar a *cell state*.

Um dos principais motivos que torna as redes neurais recorrentes tão amplamente utilizadas é a possibilidade de operação baseando-se em sequências de vetores. Essas sequências podem ser utilizadas como forma de entrada, saída ou em casos mais gerais

entradas e saídas, viabilizando a previsão de fragmentos de séries temporais do mercado FOREX.

4 Modelo Proposto de Predição

Modelos científicos fazem parte da concepção de um sistema. Estes têm o objetivo principal de simplificar sua complexidade, não deixando de reproduzir em sua essência o comportamento do sistema alvo de estudo, bem como seu entendimento. Dito isso, é apresentado abaixo o modelo utilizado como referência para este trabalho, assim como, uma versão remodelada que tem como finalidade o aperfeiçoamento do modelo base.

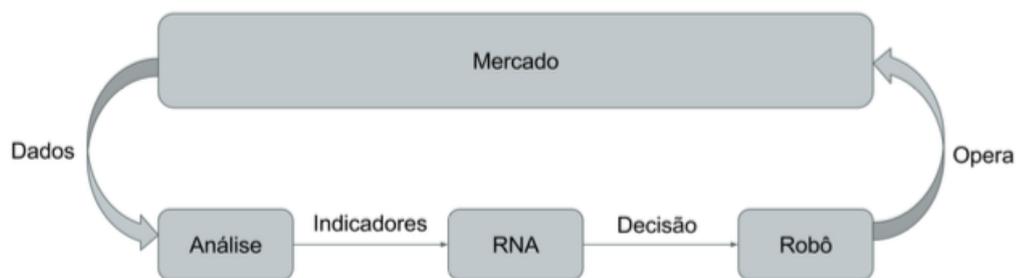


Figura 7: Modelo de previsão proposto por (Bortoli (2015))

O modelo proposto por Bortoli (2015) inicia-se com a representação do fluxo de dados originados das séries temporais do mercado FOREX. Esses dados são analisados de forma empírica, com a intenção de obter-se características do mercado financeiro, através do uso de um conjunto de dados já transformados utilizando o método de diferenças de primeira ordem, resultando assim nos indicadores usados como entrada para a rede neural. A rede neural por sua vez é encarregada pela abertura de uma ordem¹. O robô realiza operações no mercado em nome de um investidor. Ele é parametrizado e possui apenas critérios de fechamento de ordem.

A fim de melhorar tanto o desempenho computacional quanto a acurácia em relação aos resultados obtidos pela rede neural utilizada no trabalho de Bortoli (2015), optou-se pela mudança desta por uma mais apropriada do ponto de vista arquitetural, já que o problema em questão, previsão de ativos do mercado FOREX, é caracterizado por ter em seus dados propriedades sequenciais e por possuírem dependência. Parâmetros importantes como o tamanho da janela de previsão e o indicador extraído a partir dos dados do mercado FOREX também foram alterados.

Na figura abaixo pode-se observar o modelo proposto de predição com as devidas alterações.

¹ De forma análoga ao mercado financeiro, no mercado FOREX as negociações são realizadas por meio da abertura de ordens de compra e venda. Existem pelo menos três tipos de ordens segundo a definição de Garner (2012), entretanto as ordens nesse caso são as do tipo *market orders*: São abertas ao preço de mercado (*bid*) quando tratam-se de uma ordem de venda e ao preço de *ask* para compras.

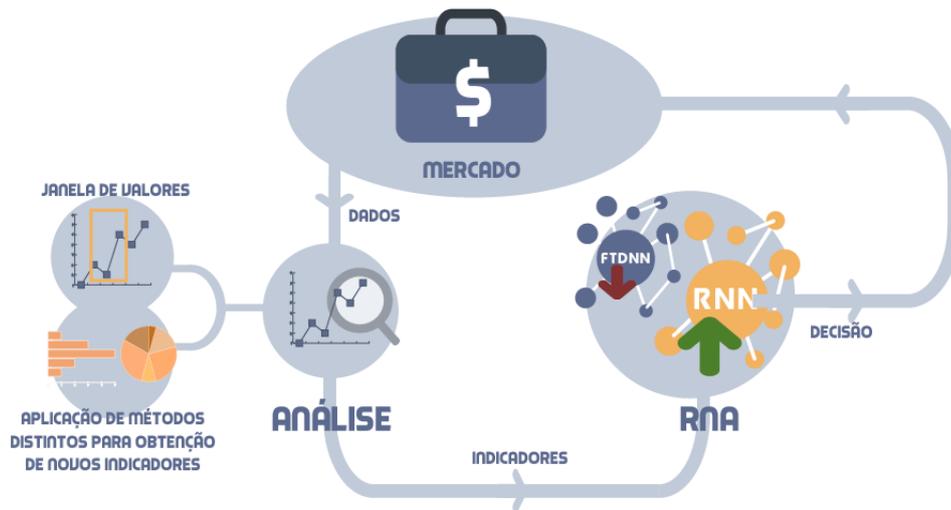


Figura 8: Modelo de previsão

O modelo proposto inicia-se com a obtenção dos dados do mercado FOREX através de uma plataforma eletrônica de negociação. Esses dados históricos são reorganizados de acordo com uma nova frequência e analisados, com o propósito de obter-se um novo indicador que possa ser utilizado como entrada para uma rede neural recorrente. Após concluída a fase de análise, a rede neural em questão recebe e processa tal informação, gerando um conjunto de dados de saída. Esses dados de saída referem-se as operações de venda e compra dentro do mercado FOREX. Estas saídas, também conhecidas como previsões, serão submetidos a comparação com um conjunto pré existente de operações, visando qualificar a taxa de acerto geral do modelo.

4.1 Metodologia

A metodologia utilizada nesse trabalho pode ser dividida em 5 partes principais apresentadas na fluxo metodológico abaixo.

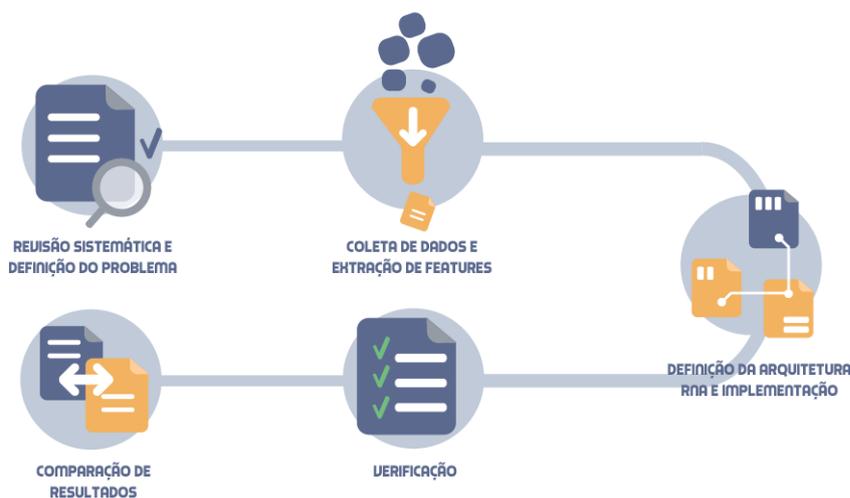


Figura 9: Fluxo metodológico

A primeira etapa do fluxo apresenta os seguintes passos:

- Elaborar a questão problema, definindo seus aspectos e limitações dentro escopo apropriado para um trabalho de conclusão de curso a nível de graduação.
- Realizar pesquisas extensas com o propósito de reunir estudos na área de inteligência artificial, predição e mercado FOREX que se assemelhem ao objetivo principal do trabalho em questão, utilizando-os como base científica.

A segunda etapa teve como objetivo a coleta de dados históricos do mercado FOREX, tendo como alvo a utilização destes como fonte para análise e criação de um novo indicador que pudesse ser utilizado neste mercado. Esta análise, foi essencialmente realizada através do uso da linguagem de programação *Python* e de bibliotecas fortemente relacionadas com a análise de dados².

A terceira etapa focou na definição e implementação, também em linguagem *python*, do que seria uma arquitetura de rede neural mais apropriada para problemas que envolvam sequências. A arquitetura escolhida foi a da rede neural recorrente, mais especificamente a *LSTM*, um tipo particular de RNN que funciona, para muitas tarefas, melhor do que a versão padrão Olah (2015).

² A biblioteca *Pandas* possui ferramentas relevantes e que simplificam o trabalho, como por exemplo: leitura e gravação de dados entre estruturas de dados em memória e em diferentes formatos, entre outras.

A quarta etapa ou etapa de verificação foi o momento de, depois de implementada e testada a rede neural, utilizar nesta o conjunto de dados anteriormente processados pela fase de análise com finalidade de obter-se as saídas necessárias para que o robô, responsável por fazer as operações no mercado, pudesse atuar.

A quinta etapa é o momento de comparar os resultados da nova abordagem proposta por este trabalho com os resultados do trabalho base, comparando os resultados das taxas efetivas de acerto e verificando se o caminho abordado aqui é mais acertado do que o anterior.

4.2 Coleta de dados e pré-processamento

A análise técnica financeira é possível devido a existência de dados históricos. Estes dados podem ser coletados de forma automatizada através de plataformas como a *MetaTrader*, também conhecida como MT4, ou até mesmo obtidos por meio de instituições que realizam operações no mercado FOREX. Juntamente com o grupo de pesquisa da Universidade do Estado da Bahia, coletou-se dados deste mercado utilizando a plataforma MT4, já que esta, além de ser gratuita e popular entre os investidores, inclui ferramentas que auxiliam na tarefa em questão, como a possibilidade de criação de *scripts* de captura de dados na linguagem de programação *MetaQuotes(MQL4)*, semelhante em sua sintaxe a linguagem C [Young \(2015\)](#).

O conjunto de dados coletados se refere ao período de um mês com a maior frequência possível, *tick-a-tick*. Um *tick* é uma medida do movimento mínimo no preço de um ativo. Neste, toda e qualquer movimentação no que diz respeito ao aumento ou diminuição no preço de um ativo é registrado, em forma de um ponto contendo um valor e o registro de tempo naquele momento. É importante frisar que nos pares de moedas considerados como de alto volume, pode-se ocorrer um número elevado de *ticks* por segundo. Tomando-se como exemplo o conjunto de dados inicial utilizado neste trabalho, verificou-se que, tanto no minuto zero quanto no segundo zero, constatou-se a existência de 70 *ticks*, isso quer dizer, 70 observações de variação no preço do par de moeda EUR/USD em 1 segundo. Com base nessa afirmação, pode-se inferir que a análise técnica referida precisa contemplar uma quantidade de registros que, no decorrer de um mês, chega a pelo menos 1 milhão de *ticks*.

4.2.1 A Análise Técnica

Com intenção de obter-se indicadores que auxiliem na predição de ativos do mercado FOREX, foi utilizado nesse trabalho a análise técnica. Durante este processo, foram realizadas observações e aplicações de conceitos matemáticos sobre os dados coletados a

fim de se chegar ao objetivo principal do trabalho. Nessa seção é apresentada a forma utilizada para extração do indicador, assim como, o indicador propriamente dito.

Para que o processo tivesse seu nível de complexidade diminuído e para que se pudesse realizar a análise de forma a manter a integridade dos dados, optou-se por redefinir o arquivo principal contendo os dados históricos *tick-a-tick* do mercado FOREX. Este foi redefinido duas vezes consecutivas. A primeira particionou o arquivo principal em arquivos menores, representando cada dia do mês. A segunda redefinição foi feita em cima destes novos arquivos, onde cada dia foi particionado levando-se em consideração suas respectivas horas, gerando assim, 24 novos arquivos. Portanto, a análise foi feita em cima desses novos conjuntos contendo uma nova frequência, hora a hora.

Para que a análise técnica pudesse ser viabilizada, fez-se necessário ainda a extração de *features* ou características destas novas séries temporais, obtendo-se assim um fragmento importante para a concepção do indicador. Essa extração pode ser abordada de duas maneiras básicas. Através da estatística descritiva ou analítica. O objetivo principal da primeira é conferir a um conjunto de dados, medidas/características geralmente descritivas. Estas podem ser classificadas como medidas de tendência central ou de variabilidade. O primeiro grupo pode incluir valores como mediana e média, enquanto o segundo pode englobar medidas como desvio padrão, variância, o valor máximo e mínimo. A estatística analítica prende-se mais a análise e comparação da natureza de dois ou mais conjuntos de dados, visando constatar a existência de diferença estatística entre eles [Vieira \(2011\)](#).

Dito isso, e fazendo-se uso da estatística descritiva, partiu-se para a definição do indicador. Este foi elaborado seguindo os seguintes passos:

- Delimitação de uma janela \mathbf{L} , tomando-se como base a frequência dos arquivos, ou seja, foram utilizadas janelas que representam uma hora cada
- Extração da *feature* média a partir dos *ticks* que formam cada série temporal
- Quantificação tanto dos *ticks* acima (TicAc) quanto abaixo (TicAb) desta média
- Aplicação da diferença entre os valores (TicAc) e (TicAb)

Empregando o conceito do componente tendência de uma série temporal, foi feita uma classificação de acordo com o resultado da diferença:

- Se $(\text{TicAc} - \text{TicAb}) > 0$ então o fragmento da série tem tendência crescente
- Se $(\text{TicAc} - \text{TicAb}) < 0$ então o fragmento da série tem tendência decrescente

- Se $(TicAc - TicAb) = 0$ então o fragmento da série não possui uma tendência suficientemente forte para nenhuma das direções

Sabendo-se o resultado da diferença, pode-se então classificar a série analisada atribuindo os valores 1, -1 e 0, seguindo respectivamente a mesma regra acima. Ao final de cada dia analisado tem-se a quantidade de 24 fragmentos de tendências, com exceção das vezes em que não se tem dados suficientes ou não foram captados registros durante todo o dia em questão. Em seguida, é realizada a soma das tendências individuais, produzindo uma tendência geral do dia analisado.

	Média dos Valores	QTD. de valores Acima da Média-TicAc	QTD. de valores abaixo da Média-TicAb	QTD. Total de cotações	Possível Tendência	Diferença
Hora 00	1.38679	511	524	1035	-1	-23
Hora 01	1.387073	304	272	576	1	32
Hora 02	1.387167	533	237	770	1	296
Hora 03	1.387047	246	140	386	1	106
Hora 04	1.387112	144	188	332	-1	-44
Hora 05	1.387174	420	681	1101	-1	-261
Hora 06	1.388434	2910	899	3809	1	2011
Hora 07	1.388767	1199	1096	2295	1	103
Hora 08	1.388574	914	1221	2135	-1	-307
Hora 09	1.388377	1184	1314	2498	-1	-130
Hora 10	1.38786	716	735	1451	-1	-19
Hora 11	1.387497	1145	1103	2248	1	42
Hora 12	1.387616	1232	1081	2313	1	151
Hora 13	1.387123	926	1314	2240	-1	-388
Hora 14	1.387081	1478	1294	2772	1	184
Hora 15	1.386584	1240	951	2191	1	289
Hora 16	1.386655	611	513	1124	1	98
Hora 17	1.386782	364	505	869	-1	-141
Hora 18	1.386789	649	1756	2405	-1	-1107
Hora 19	1.386487	1096	785	1881	1	311
Hora 20	1.386616	608	1091	1699	-1	-483
Hora 21	1.386759	387	479	866	-1	-92
Hora 22	1.386909	321	809	1130	-1	-488
Hora 23	1.386838	234	1581	1815	-1	-1347
Total do dia 01					-2	0

Figura 10: Tabela com todos valores calculados

A primeira coluna detalha as horas do dia em análise. A segunda coluna traz as médias de cada hora constituídas por quantidades diferentes de cotações. A terceira e quarta trazem o TicAc e o TicAb respectivamente. A quinta exibe a quantidade de cotações numa determinada hora. A sexta coluna é a classificação do fragmento da série, tendo como base a coluna sete, que é o resultado da diferença entre as colunas 3 e 4. A tendência do dia pode ser observada ao final da sexta coluna, tendo resultado negativo para este caso, o que significa que a série do primeiro dia tem uma tendência decrescente.

Pode-se observar na figura abaixo a série temporal analisada, plotada com todos os seus *ticks*.

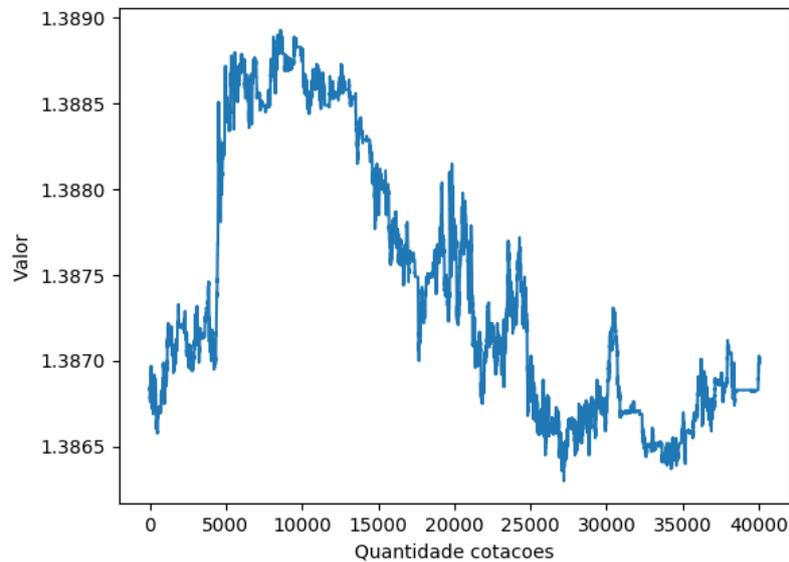


Figura 11: Série do mercado FOREX referente ao dia analisado

Abaixo pode-se observar o novo gráfico gerado a partir das médias das horas do dia. Apesar de conter substancialmente menos pontos, reproduz de forma satisfatória o comportamento geral da série.

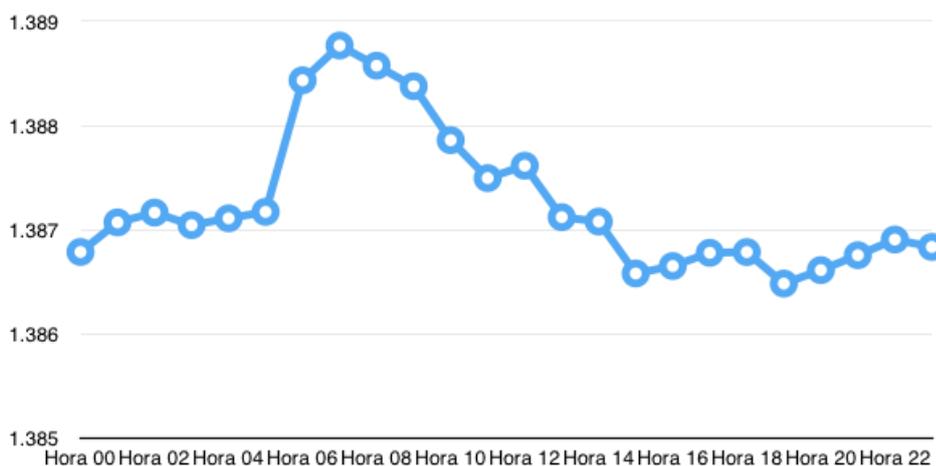


Figura 12: Médias referentes ao dia analisado

Analisando o gráfico de forma empírica, nota-se que de fato, o fragmento de série temporal estudado tem tendência decrescente, sendo representado então pelo indicador apresentado.

No decorrer da análise percebeu-se que os dias não úteis, sábados e domingos, ou continham uma quantidade de registros *ticks* insuficiente para aplicação do método pro-

posto de indicador, ou não continham registro algum. Sendo assim, decidiu-se por retirar do conjunto total de dados 9 dias. Estes são referentes ao finais de semana, acompanhando assim o funcionamento do mercado FOREX. Optou-se também por modificar a extensão do arquivo, passando de .txt para .csv (*comma-separated values*). Dentre as vantagens da mudança, destacam-se o tamanho reduzido dos seus arquivos, a relativa facilidade em ser processado por várias aplicações e por ser considerado um formato padrão.

Partindo-se desse princípio, aplicou-se o método proposto sobre todo o conjunto, buscando-se coletar dados suficientes para utilização posterior destes como entrada da rede neural.

5 A Rede Neural

Este capítulo apresenta a rede neural utilizada neste trabalho, assim como, a estratégia utilizada para seu treinamento e teste. É realizada também a análise dos resultados obtidos em relação ao desempenho da rede.

5.1 LSTM

A arquitetura de rede neural utilizada nesse trabalho foi a *Long Short Term Memory* ou *LSTM*. Durante seu desenvolvimento, tentou-se manter o menor nível de complexidade em relação a arquitetura, levando-se em consideração o problema abordado. O algoritmo de aprendizagem empregado no treinamento foi o *Adam*, usando a abordagem *Backpropagation* para ajustar os pesos da rede; para a camada escondida optou-se pela função de ativação tangente hiperbólica, também conhecida como \tanh 3.1; a função de custo aplicada ao modelo foi a Erro Quadrático Médio, ou *MSE* em inglês; para a função de ativação da camada de saída, definiu-se uma função linear, de acordo com o padrão da *API Keras Chollet et al. (2015)*.

5.2 Definição dos Hiper-parâmetros

Partindo do conceito de que, um modelo de aprendizado de máquina nada mais é do que uma fórmula matemática com um número de parâmetros que precisam ser aprendidos a partir de um conjunto de dados, hiper-parâmetros são os parâmetros que não podem ser diretamente aprendidos com o processo usual de treinamento do modelo. Portanto, esses parâmetros precisam ser atribuídos antes do início do processo de treinamento. Redes neurais podem possuir uma quantidade significativa destes, que em muitos casos são definidos de forma manual, a fim de, por exemplo, especificar a estrutura da rede.

Dito isso, pode-se observar na tabela 1 os hiper-parâmetros definidos para a rede utilizada neste trabalho, além dos já mencionados na seção 5.1.

Muito da definição desses parâmetros baseou-se no trabalho de Breuel (2015), que, através de uma investigação mais detalhada, apresentou diferentes resultados para uma série de parâmetros que podem ser empregados nas redes LSTM, visando analisar o impacto destes na performance geral do modelo.

Parâmetro	Valor
Taxa de aprendizagem do algoritmo	0.001
Tamanho do batch de treinamento	6
Número de camadas escondidas	1
Quantidade de unidades LSTM	64
Número de épocas	50
Dropout	0.4

Tabela 1: Hiper-parâmetros da rede

5.3 Preparação dos Dados para a rede LSTM

Antes de alimentar o modelo *LSTM*, é necessário que os dados utilizados sejam preparados, passando em muitos casos, por algumas transformações.

1. Transformar a série temporal para que seja possível utilizar a aprendizagem supervisionada
2. Transformar as observações para que tenham uma escala específica

O primeiro ponto é importante devido a forma com que a *API Keras* recebe os dados. É necessária a criação de um conjunto de dados além do existente, com o intuito de ser utilizado como entrada para a rede neural. Pode-se criar este novo conjunto através do deslocamento negativo de uma ou mais observações $(t-1), (t-2), \dots, (t-n)$ para cada elemento em (t) . Esta técnica é conhecida por janela deslizante.

(t-4)	(t-3)	(t-2)	(t-1)	(t)
0	0	0	0	-1
0	0	0	-1	1
0	0	-1	1	1
0	-1	1	1	1
-1	1	1	1	-1

Tabela 2: Novo conjunto de dados criado

Nota-se que a primeira linha da tabela 2 é preenchida por zeros, pois não há observação prévia para o primeiro valor em (t) que possa ser utilizado para preencher qualquer que seja a nova coluna gerada. Da mesma forma que, só a partir do quinto elemento da coluna (t) , é possível obter-se um vetor válido para utilização na rede neural, $V = [-1 \ 1 \ 1 \ 1 \ 1]$. Desse modo, aplicou-se essa técnica tanto nos dados de treinamento quanto nos de teste, possibilitando a utilização de vetores como entrada para a rede. Neste trabalho, foram utilizados vetores contendo 6 elementos cada.

Assim como em outras redes neurais, as *LSTMs* esperam que os dados estejam dentro da escala da função de ativação usada pela rede. Sabendo-se que a função de

ativação utilizada é a tangente hiperbólica 3.1, e que esta produz valores entre -1 e 1, decidiu-se então por manter o dados originais, já que após a análise técnica utilizada neste trabalho e apresentada na seção 4.2.1, os dados passam a atender tal requisito.

5.4 Estratégia de Treinamento

O treinamento de redes neurais recorrentes é conhecido por apresentar dois problemas principais: O *Exploding Gradients* e o *Vanishing Gradients*.¹ A arquitetura *LSTM* soluciona este problema, permitindo que redes maiores e mais complexas possam ser criadas. Para o treinamento da rede *LSTM*, decidiu-se por dividir o *dataset* da seguinte maneira: 70% dos dados foram separados para treinamento. Outros 15% foram reservados a fim de formar um conjunto de validação que foi utilizado durante o treinamento da rede. Os 15% restantes foram separados e destinados a criação de um conjunto para teste. Esse procedimento é conhecido como validação cruzada, sendo realizado para que o modelo preditivo possa ter sua capacidade de generalização avaliada, já que pretende-se obter valores preditos pela rede utilizando o conjunto de dados que não entrou em contato com o modelo de forma prévia, ou seja, utilizando o conjunto de teste.

A figura 13 expõe a ideia da validação cruzada, representando de forma análoga a partição dos dados descrita acima.

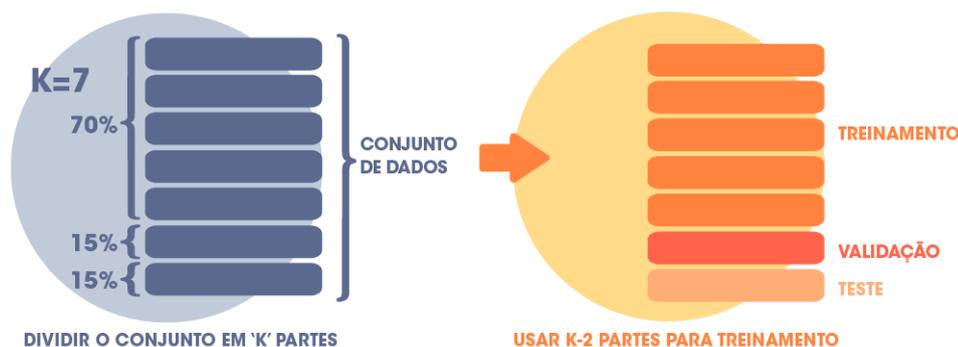


Figura 13: Divisão dos dados para validação cruzada

Depois de efetuada a divisão dos dados, foi realizada uma série de treinamentos com a finalidade de obter-se o modelo com o melhor desempenho, considerando-se que, a depender da configuração inicial da rede, resultados diversos podem ser obtidos. Para escolha da rede mais adequada, calculou-se a acurácia (acc) sobre os valores preditos pela rede durante a fase de treinamento utilizando a conjunto de validação, assim como, o erro

¹ *Exploding Gradient* refere-se ao grande aumento do gradiente durante as iterações no processo de treinamento da rede, ao passo que, o *Vanishing Gradient* refere-se justamente ao oposto. Este é um problema porquê os métodos de otimização com base em gradiente usam o valor deste para definir a direção em que os parâmetros de rede devem ser otimizados.

quadrático médio (MSE), no qual valores menores normalmente indicam um modelo mais efetivo, possibilitando uma escolha mais acertada. Dessa maneira, foram testados modelos a fim de alcançar as seguintes condições:

- A acurácia (acc) do modelo desejado deve alcançar o valor de mínimo de 0.60, quando o máximo é 1
- Erro MSE inferior à $10e^{-5}$
- Treinamentos no limite máximo de 60 épocas

Outro aspecto importante a ser considerado durante a fase de treinamento é possibilidade de superadaptação (*overfitting*). Isso acontece devido a memorização do conjunto de treinamento pelo modelo, quando o ideal, seria que este generalizasse a partir de padrões observados neste conjunto. Com a finalidade de tentar solucionar esse problema, [Srivastava et al. \(2014\)](#) propôs uma técnica de regularização de redes neurais chamada *Dropout*, que basicamente, seleciona de forma aleatória e temporária uma porcentagem pré-definida dos neurônios da rede para que, tanto suas conexões de entrada quanto de saída sejam desativadas. O efeito dessa técnica é que a rede se torna menos sensível a pesos específicos dos neurônios. Isso, por sua vez, resulta em uma rede que é capaz generalizar melhor, além de se tornar menos suscetível a superadaptação. Essa técnica foi aplicada em todas as modelos testados, variando de acordo com a topologia específica de cada modelo.

Considerando a variabilidade dos resultados da rede de acordo com a definição dos hiper-parâmetros, decidiu-se por salvar tanto o modelo quanto os pesos utilizados em cada rodada de testes, para que após a avaliação, fosse possível reproduzir os resultados.

5.5 Estratégia de Teste

Para este trabalho, testar uma rede neural significa avaliar sua capacidade de predição em relação a dados distintos daqueles utilizados durante sua fase de treinamento.

Antes da utilização dos dados na rede, aplicou-se sobre estes a análise técnica descrita na sessão 4.2.1, a fim de gerar um arquivo que pudesse ser consumido durante o treinamento e teste da rede neural. Dada a forma de funcionamento do mercado FOREX, retirou-se da análise os dias referentes aos finais de semana, devido a inconsistência da quantidade de registros coletados, restando então o equivalente a 22 dias de dados. Desses, como visto na figura 13, 15% foram separados para testar o desempenho preditivo do modelo.

Para que fosse possível haver base de comparação e para que o processo de escolha da rede mais adequada em relação ao desempenho com os dados de teste fosse simplificado,

decidiu-se então não alterar mais que 2 hiper-parâmetros para cada topologia testada durante o treinamento, visto que, treinar, ainda que seja um tipo de arquitetura, pode se tornar uma tarefa complexa caso haja uma combinação ampla de Hiper-parâmetros diferentes, dificultando assim a definição do que seria uma combinação ótima, além de em muitos casos, demandar um custo computacional alto.

Além disso, observou-se o erro (*MSE*) de forma similar ao realizado na fase de treinamento para o conjunto de teste, sabendo-se que este é comumente maior do que o valor encontrado para o conjunto de treinamento,

5.6 Interpretação das saídas da rede

Os resultados obtidos através das redes neurais podem ser difíceis de interpretar. Esta etapa é fundamental para que se possa compreender se há problemas com o modelo em questão ou até mesmo, com os dados utilizados. Neste trabalho, utilizou-se uma rede *LSTM* para prever valores que, baseando-se nos dados de entrada produzidos através da análise vista na subseção 4.2.1, deve-se girar em torno -1 e 1. Entretanto, os valores preditos pelo modelo podem estar ligeiramente abaixo de -1 ou acima de 1, fazendo-se necessária a criação de condições utilizadas para transformação desses valores de saída, com a finalidade de que estes se encaixem no padrão esperado.

Ação	Valor
Venda	-1
Compra	1
Não Operar	0

Tabela 3: Valores referência para transformação

A tabela 3 exibe os valores de referência que serão utilizados a depender da condição em que se encaixe os valores preditos pela rede neural. As condições são:

- $\text{ValorSaida} < -0,4 \rightarrow \text{Venda}$
- $\text{ValorSaida} > 0,6 \rightarrow \text{Compra}$
- $-0,4 < \text{ValorSaida} < 0,6 \rightarrow \text{Não Operar}$

6 Análise dos Resultados

Os testes para comprovação dos resultados iniciaram com o propósito de definir os melhores parâmetros para a rede neural artificial, LSTM. Para tanto, foram variadas duas características essenciais da rede: (1) quantidade de neurônios na camada oculta, (2) número de camadas ocultas. Houve ainda alteração na quantidade de épocas usada no treinamento, a fim de se encontrar um valor ideal suficiente para que a rede pudesse aprender o mecanismo gerador dos dados, sendo capaz então de realizar previsões satisfatórias. A cada uma das redes fora apresentado um universo correspondente a dezoito dias de dados exclusivamente utilizados para treinamento, dos vinte e dois disponíveis para experimentação. Para validação e quantificação da eficácia das redes, foram utilizados os quatro dias restantes.

Os resultados preliminares apontam diferentes comportamentos da rede frente a mudança de seus parâmetros, apresentando acurácia (acc) e percentual de acerto geral em relação as operações de compra e venda diferentes para cada configuração apresentada. Para calcular o percentual de acerto da rede, optou-se por seguir a forma de avaliação de desempenho proposta por [Bortoli \(2015\)](#), exibida aqui de forma sucinta:

$$Score(Compra) = \frac{N^{\circ}ComprasPrevistasCorretas}{N^{\circ}ComprasNoDataset} \quad (6.1)$$

$$Score(Venda) = \frac{N^{\circ}VendasPrevistasCorretas}{N^{\circ}VendasNoDataset} \quad (6.2)$$

Nessa maneira de avaliação, calcula-se de forma distinta o Score(Compra) e o Score(Venda). O primeiro caso é calculado através da razão entre o número de compras previstas de forma correta pela rede e a quantidade total dessa mesma ação presente no *dataset*, enquanto no segundo caso, calcula-se o número de vendas previstas corretas sobre o número de vendas total no *dataset*. Após estes cálculos, soma-se e tira-se a média dos valores encontrados, obtendo-se assim a percentual de acerto geral do modelo. Neste trabalho, são esperadas o máximo de 37 decisões de compra e 40 decisões de venda, referentes ao conjunto de dados de teste.

A tabela 4 apresenta as quatro arquiteturas de rede LSTM utilizadas para avaliação do desempenho durante a fase a treinamento. Para comparação, decidiu-se por avaliar o desempenho das redes tendo como limitador a quantidade de épocas¹. Os resultados dos

¹ Descreve o número de vezes que o algoritmo de treinamento entra em contato com todo o conjunto de dados. Então, a cada vez que o algoritmo passa pelo *dataset* de forma progressiva e regressiva, conclui-se uma época.

testes em relação a quantidade de épocas adotadas no treinamento e a acurácia do modelo podem ser encontrados nas figuras 14 e 15.

rede	camadas	neurônios
LSTM-A	1	8
LSTM-B	2	16/8
LSTM-C	1	32
LSTM-D	1	64

Tabela 4: Arquiteturas utilizadas para treinamento

O gráfico A abaixo apresenta a acurácia das redes na fase de treinamento por 30 épocas. O gráfico B exibe as mesmas redes sendo treinadas por 40 épocas. observa-se o aumento geral da acurácia dos modelos conforme o aumento das épocas de treinamento.

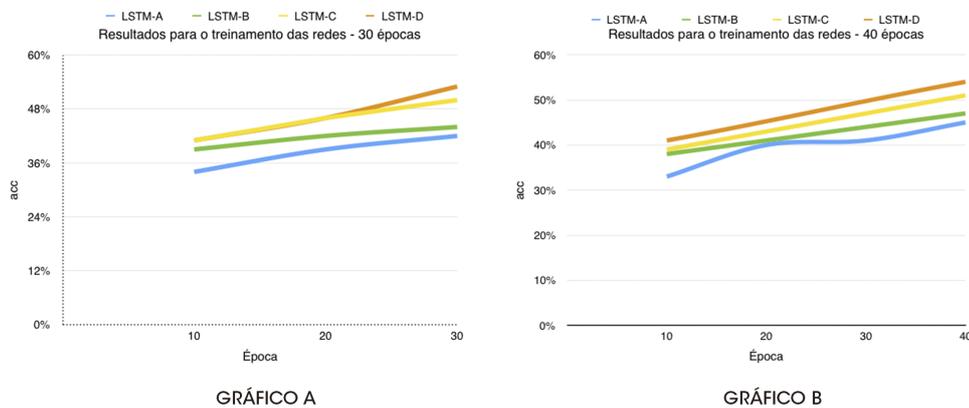


Figura 14: Resultados preliminares dos testes das redes LSTM. 30 e 40 épocas

A fim de obter-se os melhores resultados para o limite de épocas estipulado na seção 5.4, seguiu-se com os treinamento das redes aumentando de forma gradativa a quantidade de épocas, como pode-se observar abaixo.

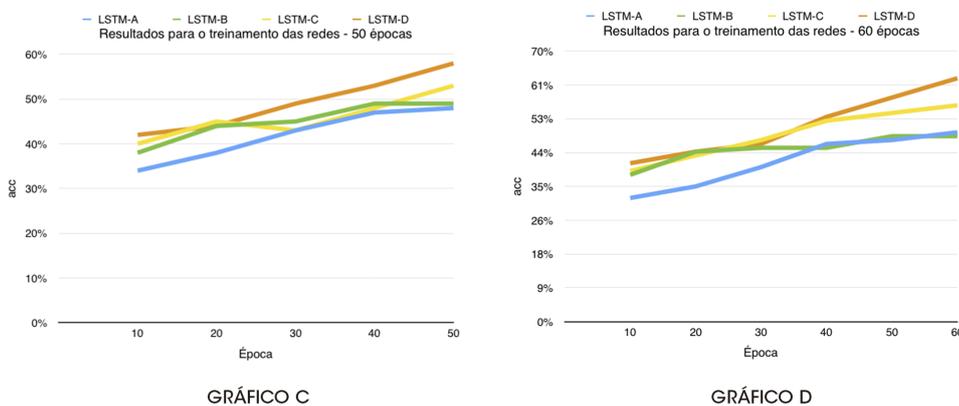


Figura 15: Resultados preliminares dos testes das redes LSTM. 50 e 60 épocas

O gráfico C da figura 15 indica que a partir da quadragésima época, as redes LSTM-A e LSTM-B apresentam uma certa estabilidade ou um aumento mínimo quanto suas respectivas acurácias (acc), sinalizando que o acréscimo de épocas pode não ser suficiente para que essas arquiteturas possam alcançar o nível de acuracidade desejado. O gráfico D comprova a hipótese mencionada, e explicita o melhor modelo entre os testados dentro da quantidade limite de épocas, a rede LSTM-D.

A partir da análise dos resultados, a melhor configuração de rede, tendo em consideração o maior acc obtido, é apresentada a tabela 5 a seguir.

rede	camadas	neurônios	acc%	épocas
LSTM-D	1	64	63	60

Tabela 5: Arquitetura de rede com melhor desempenho

Uma vez definida a rede neural, configurada e treinada, passou-se a observar o percentual de acerto das predições em função das operações de compra e venda realizadas pela rede em relação ao conjunto de teste. As colunas compra e venda referem-se as previsões da rede, ao passo que as colunas total compra e total venda referem-se as operações contidas no *dataset* de teste.

rede	compra	venda	total compra	total venda	acerto%
LSTM-D	17	25	37	40	56

Tabela 6: Taxa de acerto da rede sobre o conjunto de testes

Originalmente, a proposta da avaliação de desempenho proposta por Bortoli (2015) contempla 3 equações, sendo a última relacionada com o *score* da ação (Não Operar). Todavia, devido a natureza da análise técnica aplicada neste trabalho, não utilizou-se a terceira direção possível dos mercados, a horizontal. Esta acontece quando o mercado move-se de forma lateral, exibindo um comportamento relativamente equilibrado e tendo como ação convencional por parte dos investidores a estratégia de não operação. Dito isso, e por não haver valores preditos pela rede LSTM que contemple a condição da ação (Não Operar) apresentada na seção 5.6, retirou-se sem prejuízo para o caso específico da rede *feedforward* classificadora empregada no trabalho de Bortoli (2015), o Score(Não Operar) do resultado final, mantendo a taxa total de acerto em torno dos originais 43%.

A figura 16 exhibe os valores preditos pela rede e os valores esperados². Observa-se que a primeira barra coincide com a segunda em relação aos seus valores, indicando ambas uma tendência de queda, conseqüentemente a ação de venda. A terceira barra indica uma previsão da rede sugerindo uma permanência na tendência negativa, entretanto o resultado esperado e exibido pela quarta barra é exatamente o contrário, caracterizando-se como um erro em relação ao valor previsto pela rede.

² Cada barra de valor predito está imediatamente relacionada com a barra de valor esperado seguinte

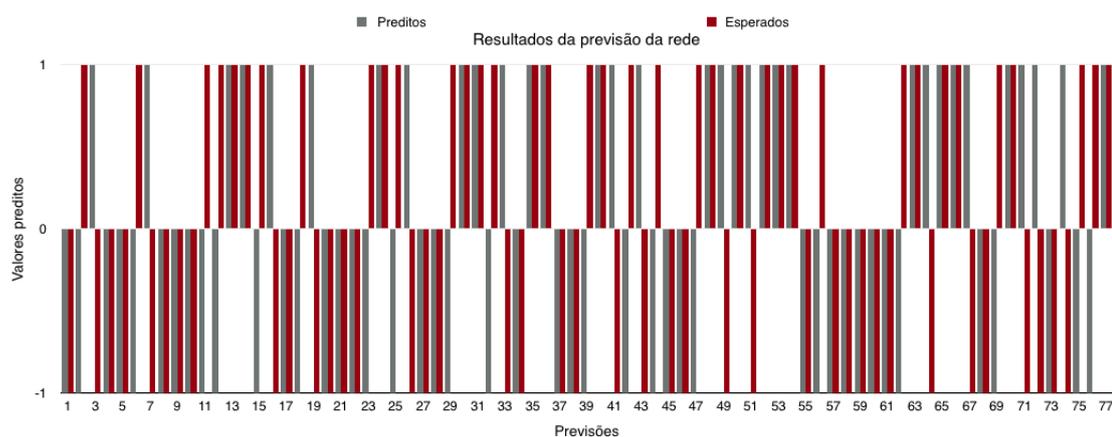


Figura 16: Valores preditos e esperados.

Verifica-se que o modelo de rede LSTM proposto nesse trabalho, quando comparado a uma rede de arquitetura *feedforward* classificadora, utilizando-se a mesma forma de avaliação de desempenho, apresentou um ganho de 13% em relação a sua taxa de previsibilidade geral, através do uso do indicador xxx.

7 Conclusões

Este trabalho teve como objetivo principal a proposição de um indicador através da análise técnica de dados reais extraídos do mercado FOREX. Esta análise tinha como um dos objetivos a criação de dados que pudessem alimentar uma rede neural, com a finalidade de prever tendências futuras neste mercado.

O indicador foi avaliado primeiramente de forma empírica a fim de se conhecer sua aplicabilidade em modelos preditivos, como modelos construídos com o auxílio das redes neurais recorrentes. Com o intuito de aprimorar a taxa de acerto, foram testadas algumas arquiteturas de rede para o problema, examinando-se particularmente a influência do aumento da quantidade de camadas, neurônios e épocas de treinamento. Entretanto, em alguns casos notou-se uma estagnação ou crescimento mínimo da acurácia de modelos com arquiteturas maiores, resultando numa performance inferior quando comparadas a redes mais simples.

Resultados preliminares demonstram claramente a eficiência do modelo preditivo proposto, chegando a apresentar eficácia de 56% na taxa de acerto. Este resultado demonstra uma melhora de aproximadamente 13% sobre a proposta apresentada por [Bortoli \(2015\)](#).

Da mesma forma, constatou-se que redes neurais do tipo LSTM podem ser significativamente melhores para tarefa de predição quando comparadas a redes classificadoras *feedforward* por exemplo, o que pode ser explicado pela capacidade intrínseca das LSTM em conservar informação relevante sobre dados passados, característica importante quando se trabalha com séries temporais. Os resultados apontam que o indicador criado pode ser utilizado para prever um dos conceitos essenciais quando se pensa na abordagem técnica para análise do mercado, a tendência. Entretanto, sabe-se que um modelo preditivo apropriado ao problema é necessário.

7.1 Discussão e Propostas

Sugere-se como ponto inicial a utilização de uma quantidade expressivamente maior de dados, tanto para treinamento quanto para teste, possibilitando assim a criação de modelos mais complexos através da aprendizagem de máquina profunda, visto que estes modelos podem apresentar uma capacidade de generalização superior.

Acredita-se ainda que ajustes mais refinados dos hiper-parâmetros da rede possam contribuir para melhora geral do desempenho desta, já que o seu desempenho está diretamente ligado ao conjunto escolhido. É possível ainda estudar a adaptação do modelo para

outras áreas do conhecimento, visto que séries temporais são encontradas numa variedade delas.

Referências

- ADHIKARI, R.; AGRAWAL, R. K. An introductory study on time series modeling and forecasting. *CoRR*, abs/1302.6613, 2013. Disponível em: <<http://arxiv.org/abs/1302-6613>>. Citado 3 vezes nas páginas 15, 17 e 18.
- ALOUD, M. et al. *Modelling the high-frequency FX market: an agent-based approach*. [S.l.], 2012. Citado na página 21.
- BORTOLI, M. F. Indicadores baseados em diferenças de primeira ordem para orientar robôs no mercado forex. 2015. Citado 9 vezes nas páginas 9, 14, 16, 21, 25, 31, 44, 46 e 48.
- BOX, G.; JENKINS, G.; REINSEL, G. *Time Series Analysis: Forecasting and Control*. Prentice Hall, 1994. (Forecasting and Control Series). ISBN 9780130607744. Disponível em: <<https://books.google.com.br/books?id=sRzvAAAAMAAJ>>. Citado 2 vezes nas páginas 15 e 25.
- BREUEL, T. M. Benchmarking of LSTM networks. *CoRR*, abs/1508.02774, 2015. Disponível em: <<http://arxiv.org/abs/1508.02774>>. Citado na página 39.
- CHOLLET, F. et al. *Keras*. [S.l.]: GitHub, 2015. <https://github.com/fchollet/keras>. Citado na página 39.
- COCHRANE, H. J. Time series for macroeconomics and finance. p. 8–9, 2005. Disponível em: <http://faculty.chicagobooth.edu/john.cochrane/research/papers/time_series_book.pdf>. Citado na página 17.
- DHAMIJA, A.; BALLA, V. Financial time series forecasting: Comparison of neural networks and arch models. *International Research Journal of Finance and Economics*, v. 1, n. 1, p. 194–197, 2010. Citado 3 vezes nas páginas 14, 15 e 25.
- DIEBOLD, F.; KILIAN, L. Measuring predictability: theory and macroeconomic applications. *Journal of Applied Econometrics*, v. 16, n. 6, p. 657–669, 2001. Disponível em: <<http://EconPapers.repec.org/RePEc:jae:japmet:v:16:y:2001:i:6:p:657-669>>. Citado na página 23.
- GUJARATI, D.; PORTER, D. *Basic Econometrics*. McGraw-Hill Irwin, 2009. (Economics series). ISBN 9780071276252. Disponível em: <<https://books.google.com.br/books?id=6l1CPgAACAAJ>>. Citado na página 19.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2nd. ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998. ISBN 0132733501. Citado 2 vezes nas páginas 25 e 26.
- HIPEL, K.; MCLEOD, A. *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier Science, 1994. (Developments in Water Science). ISBN 9780080870366. Disponível em: <<https://books.google.com.br/books?id=t1zG8OUbgdgc>>. Citado 2 vezes nas páginas 18 e 19.

- HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural Comput.*, MIT Press, Cambridge, MA, USA, v. 9, n. 8, p. 1735–1780, nov. 1997. ISSN 0899-7667. Disponível em: <<http://dx.doi.org/10.1162/neco.1997.9.8.1735>>. Citado na página 28.
- Kondratenko, V. V.; Kuperin, Y. A. Using Recurrent Neural Networks To Forecasting of Forex. *eprint arXiv:cond-mat/0304469*, abr. 2003. Citado 2 vezes nas páginas 15 e 27.
- KUNA, K. Bachelor's thesis, *Time Series prediction using neural networks*. 2015. Disponível em: <Available from WWW <http://is.muni.cz/th/410446/fi_b/>>. Citado 2 vezes nas páginas 23 e 27.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, v. 5, n. 4, p. 115–133, 1943. ISSN 1522-9602. Disponível em: <<http://dx.doi.org/10.1007/BF02478259>>. Citado na página 25.
- MURPHY, J. *Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications*. New York Institute of Finance, 1999. (New York Institute of Finance Series). ISBN 9780735200661. Disponível em: <https://books.google.com.br/books?id=5zhXEqdr_IcC>. Citado na página 24.
- OLAH, C. Understanding lstm networks. 2015. Disponível em: <<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>>. Citado 4 vezes nas páginas 9, 28, 29 e 33.
- PERSIO, L. D.; HONCHAR, O. Artificial neural networks approach to the forecast of stock market price movements. *International Journal of Economics and Management Systems*, p. 158–162, 2016. Citado na página 27.
- SHARDA, R.; PATIL, R. B. Connectionist approach to time series prediction: an empirical test. *Journal of Intelligent Manufacturing*, v. 3, n. 5, p. 317–323, 1992. ISSN 1572-8145. Disponível em: <<http://dx.doi.org/10.1007/BF01577272>>. Citado na página 15.
- SHUMWAY, R.; STOFFER, D. *Time Series Analysis and Its Applications: With R Examples*. Springer New York, 2010. (Springer Texts in Statistics). ISBN 9781441978646. Disponível em: <<https://books.google.com.br/books?id=dbS5IQ8P5gYC>>. Citado na página 14.
- SIMPSON, P. K. *Artificial Neural Systems: Foundations, Paradigms, Applications, and Implementations*. Blue Ridge Summit, PA, USA: Windcrest/McGraw-Hill, 1990. ISBN 0-07-105355-7. Citado na página 26.
- SRIVASTAVA, N. et al. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, v. 15, p. 1929–1958, 2014. Disponível em: <<http://jmlr.org/papers/v15/srivastava14a.html>>. Citado na página 42.
- VIEIRA, S. *Introdução a Bioestatística*. Elsevier Brasil, 2011. ISBN 9788535246049. Disponível em: <<https://books.google.com.br/books?id=UR7lLfnVsoC>>. Citado na página 35.
- WALL, M. Big data: Are you ready for blast-off? 2014. Disponível em: <<http://www.bbc.com/news/business-26383058>>. Citado na página 25.

- YAO, J.; TAN, C. L. A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, v. 34, p. 79–98, 2000. Citado 2 vezes nas páginas 14 e 15.
- YOUNG, A. *Expert Advisor Programming for MetaTrader 4: Creating Automated Trading Systems in the MQL4 Language*. Edgehill Publishing, 2015. ISBN 9780982645932. Disponível em: <https://books.google.com.br/books?id=_EYorgEACAAJ>. Citado na página 34.
- ZEVIANI, W. M.; Ribeiro Jr, P. J.; BONAT, W. H. *Curso em modelos de regressão não linear*. 1. ed. Campina Grande: ?, 2013. 101 p. Citado na página 27.
- ZULU, T. An introduction to the basics of forex trading. 2010. Disponível em: <<https://goo.gl/F4sHTv>>. Citado na página 22.

Apêndices

APÊNDICE A – Ambiente de desenvolvimento da rede neural utilizada

A rede neural recorrente do tipo *Long Short Term Memory* foi desenvolvida e treinada utilizando a *Application Program Interface* ou *API* de alto nível Keras, versão (2.0.5) escrita em PythonTM. Devido a utilização da Keras, a linguagem de programação utilizada foi essencialmente PythonTM, na sua versão (2.7.12), devido a sua maior estabilidade frente a versões mais recentes. A *API* supracitada foi executada sobre uma biblioteca de computação numérica chamada Theano. Utilizou-se esta por ser tratar de um projeto de código aberto, principalmente desenvolvido por um grupo de aprendizagem de máquina da Universidade de Montreal.

O trabalho foi realizado em um MacBook Pro Mid 2010, no sistema operacional OS X 10.11.6 (El Capitan), com 8gb de memória RAM e processador i5 2,53Ghz.

Pode-se ver na figura 17 a arquitetura da rede LSTM escolhida de acordo com o melhor acurácia durante o fase de treinamento, com 1 camada escondida recorrente, uma camada *Dropout* para prevenção de *overfitting*, tendo ao final uma camada *Dense* e sua ativação, responsáveis por retornarem os resultados preditos pela rede.

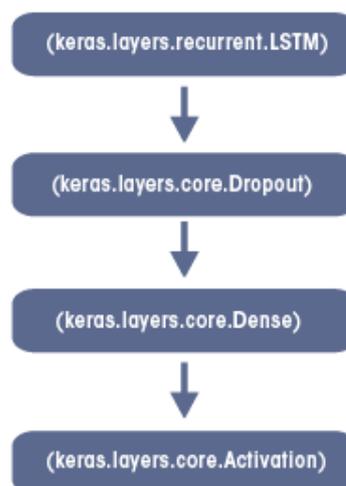


Figura 17: Arquitetura LSTM