

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DEPARTAMENTO ACADÊMICO DE COMPUTAÇÃO
CURSO DE CIÊNCIA DA COMPUTAÇÃO

PAULO HENRIQUE DOS SANTOS LUZ

**COMPARATIVO ENTRE REDE NEURAL ARTIFICIAL DE
ALIMENTAÇÃO RECORRENTE E PARA FRENTE NA PREVISÃO
DA COTAÇÃO DO CONTRATO FUTURO DO MINI ÍNDICE
BOVESPA**

TRABALHO DE CONCLUSÃO DE CURSO

MEDIANEIRA

2020

PAULO HENRIQUE DOS SANTOS LUZ

**COMPARATIVO ENTRE REDE NEURAL ARTIFICIAL DE
ALIMENTAÇÃO RECORRENTE E PARA FRENTE NA PREVISÃO
DA COTAÇÃO DO CONTRATO FUTURO DO MINI ÍNDICE
BOVESPA**

Trabalho de Conclusão de Curso apresentado ao Departamento Acadêmico de Computação da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Bacharel em Computação”.

Orientador: Prof. Dr. Paulo Lopes de Menezes

Co-orientador: Prof. Dr. Arnaldo Cândido Junior

MEDIANEIRA

2020



TERMO DE APROVAÇÃO

COMPARATIVO ENTRE REDE NEURAL ARTIFICIAL DE ALIMENTAÇÃO RECORRENTE E PARA FRENTE NA PREVISÃO DA COTAÇÃO DO CONTRATO FUTURO DO MINI ÍNDICE BOVESPA

Por

PAULO HENRIQUE DOS SANTOS LUZ

Este Trabalho de Conclusão de Curso foi apresentado às 10:30h do dia 26 de novembro de 2020 como requisito parcial para a obtenção do título de Bacharel no Curso de Ciência da Computação, da Universidade Tecnológica Federal do Paraná, Câmpus Medianeira. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Prof. Dr. Paulo Lopes de Menezes
UTFPR - Câmpus Medianeira

Prof. Pedro Luiz de Paula Filho
UTFPR - Câmpus Medianeira

Prof. Nelson Miguel Betzek
UTFPR - Câmpus Medianeira

A folha de aprovação assinada encontra-se na Coordenação do Curso.

RESUMO

Luz, Paulo Henrique dos Santos. COMPARATIVO ENTRE REDE NEURAL ARTIFICIAL DE ALIMENTAÇÃO RECORRENTE E PARA FRENTE NA PREVISÃO DA COTAÇÃO DO CONTRATO FUTURO DO MINI ÍNDICE BOVESPA. 60 f. Trabalho de Conclusão de Curso – Curso de Ciência da Computação, Universidade Tecnológica Federal do Paraná. Medianeira, 2020.

Redes neurais artificiais que contém alimentação recorrente e para frente em suas arquiteturas são ótimas ferramentas para serem utilizados em análise e predição de séries temporais, e por isso, nesta monografia foi investigado qual delas é a mais apta à ser usada em um ambiente de bolsa de valores. Foi realizado o comparativo entre a abordagem de alimentação para frente e recorrente em redes neurais artificiais sendo realizado predições por uma rede Long Short Term Memory e Multilayer Perceptron para prever a cotação do contrato futuro do mini Índice Bovespa no mês de novembro de 2019 para depois ser calculado o coeficiente de correlação de Pearson e coeficiente de determinação R^2 do conjunto de suas predições com o conjunto da cotação negociada no mês. Com os resultados dos experimentos a rede neural de alimentação recorrente obteve o maior coeficiente de correlação de Pearson quando comparada à rede de alimentação para frente e maior também que os preditores escolhidos para serem utilizados como linha de base.

Palavras-chave: APRENDIZADO DE MÁQUINA, PERCEPTRON MULTICAMADAS, REDES NEURASIS ARTIFICIAIS RECORRENTES, MERCADO FUTURO, BOLSA DE VALORES

ABSTRACT

Luz, Paulo Henrique dos Santos. COMPARISON OF A RECURRENT AND FEEDFOWARD ARTIFICIAL NEURONAL NETWORK IN THE FORECAST OF THE FUTURE OF THE MINI BOVESPA INDEX QUOTES. 60 f. Trabalho de Conclusão de Curso – Curso de Ciência da Computação, Universidade Tecnológica Federal do Paraná. Medianeira, 2020.

Artificial neural networks that use forward and recurrent feeding approach in their architectures are great tools to be used in analysis and prediction in time series, and therefore, in this monograph it was investigated which one is the most suitable to be used in a stock exchange environment. A comparison was made between the forward and recurrent feeding approach in artificial neural networks, with predictions made by a Long Short Term Memory and Multilayer Perceptron network for the quotation of the future contract of the mini Bovespa Index in November 2019, to be later calculated the Pearson correlation coefficient and the determination coefficient R^2 of the set of its predictions with the set of the price negotiated in the month. With the results of the experiments, the recurrent feeding neural network obtained the best Pearson correlation coefficient being better than the forward feeding network and also better than the predictors chosen to be used as baseline.

Keywords: MACHINE LEARNING, MULTILAYER PERCEPTRON, RECURRENT NEURAL NETWORK, FUTURES MARKETS, STOCK MARKET

”Dedico este trabalho à minha mãe Lucimara Armirio Dos Santos, que sempre esteve ao meu lado.”

”O sucesso é ir de fracasso em fracasso sem perder entusiasmo.”(Winston Churchill)

LISTA DE FIGURAS

FIGURA 1	– Mercado financeiro.	16
FIGURA 2	– Neurônio biológico e seus componentes.	20
FIGURA 3	– Neurônio artificial e seus componentes.	23
FIGURA 4	– Rede neural recorrente.	28
FIGURA 5	– Rede neural recorrente com os pesos.	29
FIGURA 6	– Rede Long Short Term Memory.	30
FIGURA 7	– Célula de estado LSTM.	31
FIGURA 8	– Portão do esquecimento LSTM.	32
FIGURA 9	– Portão de entrada LSTM.	33
FIGURA 10	– Saída da LSTM.	34
FIGURA 11	– Passos necessários para realizar o experimento.	35
FIGURA 12	– Cotação do contrato futuro do mini índice Bovespa nos conjuntos de: a) treinamento, b) validação, c) teste.	38
FIGURA 13	– Mean Square Error (MSE), Coeficiente de Pearson (R) e Coeficiente de Determinação (R ²) do período de treinamento e validação da rede de alimentação para frente.	46
FIGURA 14	– Mean Square Error (MSE), Coeficiente de Pearson (R) e Coeficiente de Determinação (R ²) do período de treinamento e validação da rede recorrente.	50
FIGURA 15	– Comparação entre as previsões das RNA avaliadas e os valores reais do conjunto de teste referente a novembro de 2019.	52
FIGURA 16	– Evolução do saldo de pontos obtidos pelo robô em dezembro de 2019. ...	54
FIGURA 17	– Predição da rede de alimentação recorrente pareada com a cotação de dezembro de 2019 do contrato futuro do mini Índice Bovespa.	55

LISTA DE TABELAS

TABELA 1	– Exemplo de 10 instâncias do domínio do conjunto de treinamento	39
TABELA 2	– Exemplo de 10 instâncias da imagem do conjunto de treinamento	40
TABELA 3	– Resultado do teste de topologias de redes neurais artificiais de alimentação para frente de 30/10/2017 à 07/12/2017	43
TABELA 4	– Resultado do teste de taxas de aprendizado de redes neurais artificiais de alimentação para frente de 30/10/2017 à 07/12/2017	44
TABELA 5	– Resultado do teste de topologias de redes neurais artificiais de alimentação recorrente de 30/10/2017 à 07/12/2017	47
TABELA 6	– Resultado do teste de taxas de aprendizado de redes neurais artificiais de alimentação recorrente de 30/10/2017 à 07/12/2017	47
TABELA 7	– Coeficientes de Pearson e R^2 obtido pelos preditores em suas predições ..	51

LISTA DE SIGLAS

API	Application Programming Interface
CSV	Comma Separated Values
ETF	Exchange Traded Fund
GRU	Gated Recurrent Unit
LSTM	Long Short Term Memory
MCP	Neurônio Artificial de McCulloch-Pitts
MLP	Multilayer Perceptron
MSE	Mean Squared Error
RNA	Rede Neural Artificial

SUMÁRIO

1 INTRODUÇÃO	11
1.1 JUSTIFICATIVA	12
1.2 TRABALHOS CORRELATOS	12
1.3 OBJETIVO GERAL E ESPECÍFICOS	14
1.3.1 Objetivos Específicos	14
2 REFERENCIAL TEÓRICO	15
2.1 MERCADO FINANCEIRO	15
2.1.1 Bolsa de Valores	16
2.1.2 Índice Bovespa	17
2.1.3 Mercado de Contratos Derivativos	17
2.1.4 Mercado de Contratos Futuros	18
2.1.5 Contrato Futuro do Mini Índice Bovespa	19
2.2 REDES NEURAIS ARTIFICIAIS	19
2.2.1 Tecido Nervoso	19
2.2.2 História das Redes Neurais Artificiais	20
2.2.3 Teoria Hebbiana	21
2.2.4 Primeiras Redes Neurais Artificiais	21
2.2.5 O Ressurgimento do Interesse da Comunidade Científica em Rna	22
2.2.6 Componentes das Redes Neurais Artificiais	22
2.2.7 Neurônio Artificial	23
2.2.8 Funções de Ativação	24
2.2.9 Arquitetura, Disposição de Neurônios entre Camadas e Fluxo de Informações	25
2.2.10 Aprendizado das Redes Neurais Artificiais	25
2.2.11 Algoritmo de Treinamento <i>Backpropagation</i>	26
2.2.12 Perceptron Multicamadas	27
2.2.13 Redes Neurais Artificiais Recorrentes	28
2.2.14 A RNA <i>Long Short Term Memory</i>	30
3 MATERIAIS E MÉTODOS	35
3.1 CONFIGURAR AMBIENTE DE EXPERIMENTOS	35
3.2 ORGANIZAR FERRAMENTAS DE DESENVOLVIMENTO	36
3.3 COLETAR BASE DE DADOS	37
3.4 ESCOLHER MÉTRICAS PARA AVALIAR O DESEMPENHO DAS RNAS	40
3.5 PROJETAR E IMPLEMENTAR RNA DE ALIMENTAÇÃO RECORRENTE E ADIANTE	41
3.6 AVALIAR A PREDIÇÃO DOS MODELOS NEURAIS	42
4 RESULTADOS OBTIDOS E DISCUSSÃO	43
4.1 ELABORAÇÃO E CONCEPÇÃO DO MODELO NÃO RECORRENTE	43
4.2 OTIMIZAÇÃO DO MODELO DE ALIMENTAÇÃO PARA FRENTE	44
4.3 ELABORAÇÃO E CONCEPÇÃO DO MODELO RECORRENTE	46
4.4 OTIMIZAÇÃO DO MODELO RECORRENTE	48
4.5 ANÁLISE DAS PREDIÇÕES	51

4.6	SIMULAÇÃO DE NEGOCIAÇÃO NO MERCADO DE CONTRATO FUTURO DO MINI ÍNDICE BOVESPA EM DEZEMBRO DE 2019	53
5	CONCLUSÕES	56
5.1	TRABALHOS FUTUROS	57
	REFERÊNCIAS	58

1 INTRODUÇÃO

Aprendizado de máquina é a técnica que ensina o computador a executar uma tarefa por meio da repetitiva apresentação de dados de como deve ser feito a sua solução. Redes neurais artificiais são modelos computacionais que fazem uso do aprendizado de máquina e vem obtendo destaque por resolverem problemas complicados de serem programados, sendo capazes de abstrair a solução de problemas pela apresentação de dados relacionados ao seu contexto e solução aos neurônios artificiais.

Classificadores, modelos regressivos e preditivos que são criados usando Redes Neurais Artificiais são muito eficazes, o que os tornam amplamente usados em áreas como visão computacional, reconhecimento de som, tradução automática e ainda há um extenso campo para serem aplicados, inclusive na busca por lucro em bolsas de valores.

Ao arquitetar uma Rede Neural Artificial (RNA) é necessário pensar primeiro em que tipo de problema é preciso resolver, por exemplo, conexões recorrentes podem não ter nenhuma utilidade ao ser aplicadas em classificação de imagem mas podem fazer muita diferença quando o objetivo é classificar um vídeo.

Mas ainda é difícil saber quais benefícios são inseridos ao adotar um padrão de conexão específica em um modelo de RNA, pois uma RNA pode diferenciar-se de outra por diversas características, podendo variar, por exemplo, em quantidades de camadas e neurônios ou até mesmo funções de ativações diferentes entre camadas.

Na bolsa de valores os investidores utilizam maioritariamente dois métodos para prever a tendência de um ativo sendo o primeiro método a análise fundamentalista e o segundo método a análise técnica do preço. A análise fundamentalista consiste em verificar os resultados financeiros da empresa e a análise técnica busca encontrar padrões recorrentes no gráfico para poder diagnosticar a direção da tendência de seu preço.

Como existe uma grande quantidade de dados armazenados em forma de séries temporais sobre ativos negociados em bolsas de valores é plausível a possibilidade de o investidor delegar a tarefa de encontrar ações com potencial de lucro para redes neurais artificiais, entretanto as abordagens de rede neurais artificiais que usam de alimentação recorrente e para frente são boas ferramentas de predição e análise de séries temporais, no

que pode levar o investidor à ter dúvidas na escolha de qual abordagem utilizar caso queira automatizar suas negociações na bolsa de valores com redes neurais artificiais.

Considerando isto, a presente monografia faz um comparativo de predições no preço do contrato futuro do mini Índice Bovespa de uma RNA que faz uso de conexões recorrentes com uma outra de alimentação para frente utilizando o auxílio das métricas de correlações estatísticas.

1.1 JUSTIFICATIVA

Na bolsa de valores se junta psicologia comportamental com grandes volumes de dados, é um local no qual padrões comportamentais influenciam a movimentação do preço toda vez que uma nova informação é publicada, um campo difícil de desenvolver soluções por meio da análise de especialistas mas que pode se beneficiar muito de soluções que o aprendizado de máquina pode proporcionar.

Se o investidor quiser começar a fazer uso de modelos regressivos baseado em RNA para prever a tendência do preço um dos passos é entender as vantagens e desvantagens que o uso de cada arquitetura de RNA pode trazer à sua ferramenta e principalmente, devido a natureza temporal dos preços, conhecer quais são os efeitos que conexões recorrentes e de alimentação para frente podem trazer ao modelo preditivo.

Portanto fazer um comparativo entre alimentação recorrente e para frente no mercado de contratos futuros do mini Índice Bovespa pode munir o investidor com informações valiosas para quando ele começar a utilizar RNA como ferramenta em seus investimentos o que pode resultar em mais eficiência na bolsa de valores como um todo e assim melhorar a alocação de recursos na economia.

1.2 TRABALHOS CORRELATOS

Existem diversos estudos de modelagem de RNA para predição de cotação de ativos em bolsas de valores bem como estudos de modelagem de RNA com arquiteturas diferentes

para comparar a acurácia de suas previsões.

Schossler (2017), modelou uma rede *Multilayer Perceptron* (MLP) adaptada a receber séries temporais para ajudar o investidor na compra e venda de ações de curto prazo na bolsa de valores. Schossler (2017) modelou uma rede neural artificial de alimentação para frente que treinou por 5000 épocas recebendo na camada de entrada informações da cotação de abertura, máxima, mínima e fechamento e volume de negociação em velas de 15 minutos para prever o valor de mercado das ações da empresa americana Tesla Motors negociada na NASDAQ conseguindo obter o valor de 0.1203 para a métrica *Mean Absolute Error* e testou seu modelo de predição no dia 15/02/2017 obtendo um valor financeiro de \$10.3444.

Nelson (2017) modelou uma rede *Long Short Term Memory* (LSTM) que utiliza o histórico de preço junto com indicadores de análise técnica para prever tendências de preços de ações em um curto prazo. A LSTM criada por Nelson (2017) foi considerado o modelo mais vantajoso para a análise de séries temporais quando comparado aos preditores usados como linha de base.

Giacomel (2016) propôs um método de compra e venda de ações para investidores de perfil moderado, e um método para investidores de perfil agressivo utilizando-se de *ensembles* de RNA. O *ensemble* para perfil de investidor moderado se mostrou um método com lucros consistentes enquanto o *ensemble* de perfil agressivo fez as melhores previsões para as séries temporais e pode ser indicado para ser usado como uma ferramenta na estratégia de um investidor arrojado.

Kuniaki et al. (2016) utilizou-se de modelos de *deep learning*, vetor de parágrafo e LSTM para fazer previsões em bolsas de valores utilizando de diferentes fatores como notícias, histórico de preço, diferentes índices e variáveis econômicas. Os resultados experimentais do Kuniaki et al. (2016) indicaram que os modelos que consumiam entradas de dados textuais foram melhores do que os que apenas usaram dados numéricos e entre os modelos preditivos a LSTM foi o modelo que melhor conseguiu captar a influência dos dados textuais nas séries temporais.

Adebiyi et al. (2016) propôs um modelo híbrido de uma rede MLP com analisadores técnicos para predizer o futuro do preço das ações, obtendo melhores resultados do que utilizando apenas indicadores de análise técnica.

Sundermeyer et al. (2015) comparou modelos de contagem e RNA *feedforward*, recorrente clássicas e LSTM na tarefa de reconhecimento de fala, avaliando os modelos em perplexidade e taxa de erro de palavras, obtendo melhores resultados nas RNAs, com uma clara vantagem das redes LSTM sobre outras RNAs.

Chung et al. (2014) comparou uma RNA recorrente clássica, uma LSTM, e uma *Gated*

Recurrent Unit (GRU) na modelagem de sinal de fala e de música polifônica, obtendo clara vantagem das LSTM e GRU sobre a rede recorrente clássica, e uma pequena vantagem da rede GRU sobre a LSTM.

1.3 OBJETIVO GERAL E ESPECÍFICOS

O objetivo geral deste trabalho é comparar os conjuntos de predições obtidos por meio da utilização de RNA com alimentação recorrente em suas camadas e de RNA padrão que só usa conexões para frente no mercado de contrato futuro do mini Índice Bovespa no mês de novembro de 2019.

1.3.1 Objetivos Específicos

- Definir quais métricas estatísticas serão utilizadas para analisar as predições;
- Criar a base de dados que será utilizada na otimização das RNAs;
- Projetar RNA de alimentação recorrente e para frente que realizará predições em novembro de 2019;
- Escolher dois preditores arbitrariamente para que suas predições sejam usadas como linha de base;
- Analisar os resultados obtidos das predições em novembro de 2019 e realizar um comparativo das RNAs com os preditores linha de base levando em consideração o quão relacionado estão suas predições com a realidade.

2 REFERENCIAL TEÓRICO

Este capítulo tem como propósito discorrer sobre o referencial teórico desta monografia, nele é apresentado na primeira seção o sistema financeiro com foco no mercado de renda variável apresentando o funcionamento do mercado de contratos futuros do mini Índice Bovespa que é negociado na bolsa de valores. Na segunda seção foi introduzido ao leitor conhecimentos sobre redes neurais artificiais que segundo Goyal et al. (2018) se deve ao crescimento do poder computacional das unidades gráficas de processamento e ao fácil acesso a cada vez mais e maiores conjuntos de dados pela Internet permitiu ao aprendizado de máquina baseado em RNAs se tornar cada vez mais predominante em atividades que pessoas executam bem, possibilitando que computador realize atividades como reconhecimento de imagens, reconhecimento de escrita e fala pelo aprendizado obtido a partir da observação de dados.

2.1 MERCADO FINANCEIRO

O mercado financeiro tem como objetivo conectar agentes superavitários com agentes deficitários. Os agentes superavitários são pessoas ou empresas que possuem poupança enquanto os agentes deficitários são pessoas ou empresas que consomem mais do que são capazes de produzir. Para atender as necessidades de seus participantes o mercado financeiro se divide em quatro grandes mercados: Mercado de crédito, mercado de capitais, mercado monetário e mercado de câmbio, como retratado na Figura 1 (FORTUNA, 1992; FISCHER et al., 2014; MENESES; MARIANO, 2011).

No mercado de câmbio se negocia a troca de moedas estrangeiras por moeda nacional. No mercado de capitais ocorre a transferência de recursos entre agentes superavitários e deficitários por meio das negociações de valores mobiliários. No mercado monetário se negocia títulos de moeda entre o banco central e instituições financeiras autorizadas e por



Figura 1 – Mercado financeiro.

Fonte: (FISCHER et al., 2014)

fim, no mercado de crédito acontece negociações de empréstimo entre agentes superavitários e deficitários por intermédio de uma instituição financeira (TOSCANO, 2004; FISCHER et al., 2014; FORTUNA, 1992).

2.1.1 Bolsa de Valores

A bolsa de valores surgiu da necessidade de criar um mercado secundário para fornecer liquidez e estabelecer um ambiente organizado, transparente e propício ao livre mercado que fornece confiança para o investidor caso precise liquidar ou compensar um investimento em valores mobiliários (FISCHER et al., 2014; FORTUNA, 1992).

Valores mobiliários são contratos que organizações privadas utilizam para arrecadar recursos diretamente dos investidores pois estão definidos nele os deveres e direitos dos participantes incluindo a possibilidade de sua posse ser transferida (FISCHER et al., 2014; FORTUNA, 1992; TOSCANO, 2004).

A precificação na bolsa é realizada pela oferta e da demanda dos investidores pelos

valores imobiliários. Os efeitos de sua precificação refletem a situação econômica do país ou do setor em que o ativo está relacionado. No Brasil a bolsa de valores em atividade é a B3, uma sociedade de capital aberto com o propósito de criar e gerenciar sistemas de negociações, compensação, liquidação, depósito e registro de ativos, além de atuar como contraparte garantidora de crédito para ativos listados (B3, 2018; FORTUNA, 1992).

2.1.2 Índice Bovespa

O Índice Bovespa foi criado em 1968 pela Bolsa de Valores de São Paulo (Bovespa) e mantido atualmente pela B3. Este índice é uma carteira de ações teórica com a função de indicar o desempenho médio dos ativos de maior negociabilidade e representatividade no mercado de ações brasileiro sendo sua composição reajustada a cada 4 meses (FISCHER et al., 2014; MENESES; MARIANO, 2011).

Toscano (2004) incrementa que o Índice Bovespa é um índice genérico, ou seja, existe a possibilidade de haver ações pertencentes ao índice que estejam passando por uma queda no preço enquanto o Índice Bovespa passa por uma alta na sua cotação e vice e versa.

2.1.3 Mercado de Contratos Derivativos

Mercado de derivativos é o nome que se dá ao conjunto de mercados em que se negocia contratos que tem seu valor derivado de um ativo físico ou financeiro. O contrato derivativo é um valor mobiliário no qual contém especificações sobre um ativo tácito e sua negociação permite a transferência de risco entre agentes financeiros (FISCHER et al., 2014; MENESES; MARIANO, 2011).

Os contratos derivativos podem ser classificados em contrato a termo no qual o comprador e vendedor entram em acordo no presente sobre uma negociação que se dará no futuro, com direitos e deveres das contrapartes definidas em um contrato ou customizado ou padronizado ou podem ser classificados de exercícios contingenciais onde os pagamentos são realizados apenas depois de as condições predeterminadas forem atingidas (ASSAF, 2014).

Este conjunto de mercados dão a oportunidade que seus participantes possam proteger seus ativos da flutuação de preços realizando uma operação de *hedge*, na qual o agente

econômico possui o ativo e vende um contrato derivado deste ativo na bolsa de valores para se proteger da queda do seu preço ou, por outro lado, caso o agente precise adquirir um ativo ou travar preço de uma dívida estrangeira à ser paga num futuro próximo ele realiza uma operação de compra do contrato derivado do ativo no mercado de derivativos para se proteger da alta do preço do ativo (TOSCANO, 2004; MENESES; MARIANO, 2011).

Além de se proteger da flutuação do preço, também é possível alavancar o capital usando-se do mercado de contratos derivativos. Neste caso o agente econômico atua como especulador pois ele não possui o ativo mas realiza compras e vendas no mercado de acordo com suas expectativas na movimentação de seu preço (TOSCANO, 2004; MENESES; MARIANO, 2011).

Fischer et al. (2014), Meneses e Mariano (2011) comentam sobre um terceiro caso de agente econômico que tem como objetivo obter lucro realizando operações de arbitragem tirando vantagem da diferença de preço de um mesmo produto negociado em bolsas diferentes.

2.1.4 Mercado de Contratos Futuros

O mercado de contratos futuros é negociado apenas em ambiente de bolsas de valores, neste mercado quando se consolida uma negociação as contrapartes concordam em manter o poder de compra e de venda sobre um ativo desde o dia em que foi negociado até o prazo de vencimento, devido a isto o mercado de contratos futuros é classificado como um mercado de contrato a termo dentro dos mercados de contratos derivativos como definido na seção 2.1 (TOSCANO, 2004; FISCHER et al., 2014).

A quitação do contrato dos participantes envolvidos na negociação é realizada pela câmara de compensação no qual é a solução encontrada pelos corretores para diminuir o risco de crédito na data de liquidação de contrato dos participantes dos diversos mercados de derivativos (ASSAF, 2014; MENESES; MARIANO, 2011; FORTUNA, 1992; MENDEL et al., 2006).

Para reduzir o risco de crédito a câmara de compensação obriga todos os participantes a depositarem um montante de dinheiro predefinido pela bolsa como margem de garantia para que o negociante possa encarregar-se com os ajustes diários. O ajuste diário é uma ação de crédito ou débito de acordo com o último preço fechado no final da sessão de negociação na margem de garantia dos negociantes envolvidos (ASSAF, 2014; MENESES; MARIANO, 2011; FORTUNA, 1992; MENDEL et al., 2006).

Caso a margem de garantia de uma das partes da negociação esteja abaixo do valor

estipulado pela câmara de compensação devido a um débito no último ajuste diário e não haja o reabastecimento da margem por sua parte então a sua posição no contrato será finalizada com a primeira e melhor oferta no leilão da próxima abertura de negociação (ASSAF, 2014; MENESES; MARIANO, 2011; FORTUNA, 1992; MENDEL et al., 2006).

2.1.5 Contrato Futuro do Mini Índice Bovespa

É chamado mini Índice Bovespa o montante de 10% do Índice Bovespa. A B3 organiza e mantém um sistema de negociação para contrato futuro do mini índice Bovespa, o principal objetivo deste mercado é facilitar o ingresso das pessoas físicas e de pequenas empresas que o usa para proteger suas ações ou *Exchange Traded Funds* (ETF), da oscilação dos preços, ou para assumir riscos sobre seu capital utilizando da especulação (B3, 2018; MENESES; MARIANO, 2011).

2.2 REDES NEURAIS ARTIFICIAIS

Redes neurais artificiais são modelos computacionais que usam uma arquitetura de processadores distribuídos paralelamente em camadas conectadas por inúmeras conexões ponderadas sendo que cada unidade de processamento executa uma função matemática para realizar decisões simples (BRAGA et al., 2000).

Uma RNA é um modelo abstrato que tem como inspiração o funcionamento do cérebro, fruto da necessidade de criar máquinas capazes de aprender e generalizar além de adaptar-se ao ambiente. Os estudos sobre o tema buscam entender como funciona o cérebro humano e se possível replicá-lo artificialmente (FACELI et al., 2011; HAYKIN, 2001).

2.2.1 Tecido Nervoso

O tecido nervoso está distribuído no corpo humano e forma uma rede de comunicação

tal qual é possível classificá-lo em sistema nervoso central e sistema nervoso periférico. Os principais componentes do tecido nervoso é o neurônio e a célula de glia sendo o neurônio a célula que recebe e passa estímulos nervosos enquanto a célula de glia tem a função de dar suporte aos neurônios (JUNQUEIRA; CARNEIRO, 2008).

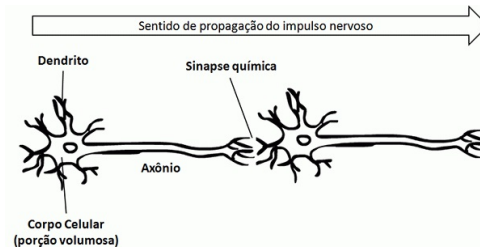


Figura 2 – Neurônio biológico e seus componentes.

Fonte: (MENDES, 2013)

As informações recebidas pelos órgãos sensoriais são conduzidas pelos nervos até o encéfalo onde são processadas por uma rede de neurônios que produz uma resposta para a situação. Os neurônios (Figura 2) são formados por um corpo celular de onde saem diversas ramificações chamadas de dendritos que são especializadas em receber sinais nervosos de outros neurônios ou células sensoriais e também de um único filamento, denominado axônio, para passar o sinal nervoso recebido pelos dendritos (LINHARES; GEWANDSZNAJDER, 2006).

O lugar onde ocorre a ligação das ramificações entre dois neurônios é chamado de sinapse, geralmente os sinais são repassados do axônio para os dendritos por uma região em que essas ramificações ficam lado a lado, quando o impulso chega ao axônio são liberadas substâncias químicas que auxiliam a transmissão e são eliminadas logo depois de o impulso ser conduzido (SILVA; SASSOM, 2005).

2.2.2 História das Redes Neurais Artificiais

Esta seção trata da evolução das RNAs ao longo dos anos. Entre 1943 e 1958 diversos conceitos utilizados em RNAs foram formulados, por exemplo, McCulloch e Pitts modelaram em 1943 o neurônio artificial de McCulloch-Pitts (MCP), Neste estudo foi proposto o neurônio MCP que pode realizar funções lógicas simples em que a combinação de vários MCP resulta em um sistema de elevado poder computacional.

É importante observar que o neurônio MCP modela o poder computacional do

neurônio biológico, mas não reproduz seu aprendizado tal que seu estudo foi iniciado por Donald Hebb em 1949. Mais adiante, em 1958, Rosenblatt propôs a rede Perceptron que usa o neurônio artificial de McCulloch e Pitts com um algoritmo de aprendizado (HAYKIN, 2001).

2.2.3 Teoria Hebbiana

Hebb (1949) postulou que neurônios que disparam sinais nervosos ao mesmo tempo tendem a criar conexões mais fortes enquanto os que disparam assincronamente tendem a ter suas conexões enfraquecidas, ou seja, se o neurônio pré-sináptico tem grande influência na ativação do neurônio pós-sináptico a ligação entre os dois tende a ganhar maior relevância enquanto se o neurônio pré-sináptico tem pouca influência na ativação do neurônio pós-sináptico a ligação entre os dois tende a enfraquecer (BRAGA et al., 2000).

Ao aplicar o aprendizado hebbiano nas ligações entre neurônios artificiais resulta que se dois neurônios são ativados juntos, o valor do peso de sua conexão aumenta e quando os dois neurônios disparam separados o valor do peso de sua conexão diminui. Usando a lei de Hebb, uma rede neural é capaz de aprender a associar disparos de neurônios (COPPIN, 2010).

2.2.4 Primeiras Redes Neurais Artificiais

O Perceptron foi a primeira RNA publicada, seu criador foi o Rosenblatt, que, em 1958 utilizou o modelo de neurônio MCP com o primeiro método de treinamento para RNAs. O Perceptron é alimentado por entradas por intermédio de uma máscara de pré-processamento e fornece uma saída binária como saída da rede (FACELI et al., 2011).

Com seu trabalho Rosenblatt trouxe um novo ponto de vista para o reconhecimento de padrões com aprendizagem supervisionada em que o grande diferencial de seu trabalho foi a demonstração do teorema de convergência do Perceptron no que consta que se um conjunto de entradas pode-se separar instâncias de duas classes diferentes através de uma reta no gráfico bidimensional, o Perceptron pode classificar a qual classe cada entrada pertence (HAYKIN, 2001; COPPIN, 2010).

Widrow e Hoff (1960) publicaram a rede ADALINE, uma rede similar ao Perceptron e diferente apenas em seus algoritmos de treinamentos tal que na ADALINE usa-se o algoritmo

denominado regra delta que adapta os pesos da rede em função do erro de sua saída linear antes de aplicar a função de ativação. A ADALINE, assim como o Perceptron, consegue aprender e modelar diversas funções, porém limitadas a funções linearmente separáveis (COPPIN, 2010; BRAGA et al., 2000).

Minsky e Papert (1969) provaram em seu livro as limitações do Perceptron que era incapaz de fazer generalizações globais usando aprendizado local e esse problema persistia para outras variantes do modelo descrito por Rosenblatt, assim como para Perceptrons com múltiplas camadas (HAYKIN, 2001).

2.2.5 O Ressurgimento do Interesse da Comunidade Científica em Rna

Houve pouco avanço na década de 1970, o pessimismo gerado pelo trabalho de Minsky e Papert (1969) acerca das limitações do Perceptron ocasionou em dificuldades em captar os recursos necessários para acesso a máquinas capazes de realizar experimentos, o que desencorajou muitas pesquisas sobre RNA (HAYKIN, 2001).

Na década 1980, as pesquisas sobre RNA ressurgem fortemente com HOPFIELD (1982), apontando propriedades associativas das RNAs e relações das redes recorrentes auto associativas com sistemas físicos e também pela descrição do algoritmo *Backpropagation*, o qual revela que Minks e Papert foram pessimistas demais em seu estudo. Outro motivo que ajudou na retomada das pesquisas sobre RNA foi o avanço da tecnologia e o fracasso de outros modelos de aprendizagem de máquina em realizar tarefas simples para humanos (BRAGA et al., 2000).

2.2.6 Componentes das Redes Neurais Artificiais

Esta seção busca conceituar e explicar os componentes de uma RNA, será discutido o neurônio artificial Perceptron, funções de ativação, assim como arquiteturas, padrões de conexões fluxo de informações pela rede e, por fim, discutir as maneiras de fazer a rede aprender, incluindo o detalhamento do *Backpropagation*.

2.2.7 Neurônio Artificial

A Figura 3 representa o neurônio Perceptron nomeado k . Suas entradas, pesos e *biases* são definidos respectivamente como x_i , w_{ki} e b_k e o potencial de ativação v_k sendo definido na Equação 2.1:

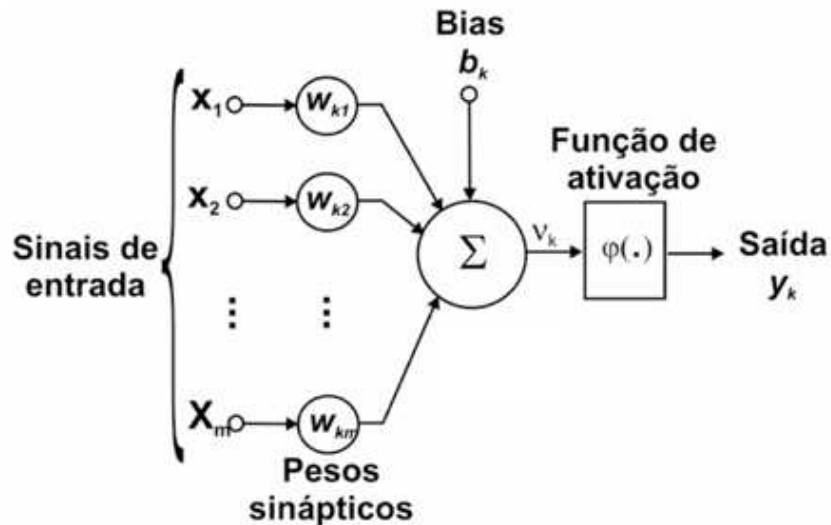


Figura 3 – Neurônio artificial e seus componentes.

Fonte: (ZAMBIASI, 2018)

$$v_k = \sum_{i=1}^m x_{ki} w_{ki} + b_k \quad (2.1)$$

Onde:

v_k = potencial de ativação do neurônio k ;

x_{km} = m ésima entrada do neurônio k ;

w_{km} = peso da m ésima conexão sináptica do neurônio k ;

b_k = bias do neurônio k .

A adição de b_k é equivalente a adicionar uma conexão com entrada fixa de valor 1 multiplicada por um peso w_k resultando na Equação 2.2:

$$v_k = \sum_{i=1}^m x_{ki} w_{ki} \quad (2.2)$$

A função de ativação $\varphi(v_k)$ usa o potencial de ativação v_k para gerar a saída y_k da rede como definida na Equação 2.3:

$$\varphi(v_k) = y_k \quad (2.3)$$

Sendo que:

φ = é a função de ativação utilizada pelo neurônio k ;

v_k = potencial de ativação obtido pelo neurônio k ;

y_k = é a resposta de saída processada pela função de ativação do neurônio k .

(HAYKIN, 2001).

2.2.8 Funções de Ativação

Como visto na seção do neurônio artificial e reforçado por Coppin (2010), o neurônio artificial decide se irá produzir um sinal usando uma função de ativação $\varphi(v_k)$. Haykin (2001) identifica três tipos de funções de ativação básicas: Função de limiar, função linear por partes e função sigmoide logística.

A função limiar (degrau ou passo), na Equação (2.4), tem como saída um valor $\varphi(v_k) = 1$ caso v_k na função seja maior ou igual a 0 ou terá um valor para $\varphi(v_k) = 0$ caso ocorra a entrada v_k na função seja menor que 0 (BRAGA et al., 2000).

$$\varphi(v_k) = \begin{cases} 1, & v_k \geq 0 \\ 0, & v_k < 0 \end{cases} \quad (2.4)$$

Na Equação (2.5) está definida uma função de ativação linear por partes, ela é semelhante a função degrau. A função linear por partes será $\varphi(v_k) = v_k$ se $1 > v_k > 0$ então a saída será a entrada v_k que está entre 0 e 1 (HAYKIN, 2001).

$$\varphi(v_k) = \begin{cases} 1, & v_k \geq 1 \\ v_k, & 1 > v_k > 0 \\ 0, & v_k \leq 0 \end{cases} \quad (2.5)$$

Haykin (2001) aponta que função de ativação sigmoide logística tem intervalo contínuo de 0 a 1 e é diferenciável, possui formato de “s”, é a mais comum das funções de ativações em RNAs. Faceli et al. (2011) ainda acrescentam que a sigmoide é uma representação diferenciável da função limiar e pode ter variadas inclinações u . A Equação (2.6) representa a função sigmoide logística.

$$\varphi(v_k) = \frac{1}{1 + e^{-(uv_k)}} \quad (2.6)$$

No que:

u = ângulo de inclinação k . neurônio k ;

2.2.9 Arquitetura, Disposição de Neurônios entre Camadas e Fluxo de Informações

Em termos de arquitetura, uma RNA pode ter inúmeros ou apenas um neurônio em uma camada consistindo de entrada e saída, sendo reconhecidas como redes de uma camada, ou diversos neurônios entre as camadas de entrada e saída da rede sendo conhecidos como redes de múltiplas camadas (BRAGA et al., 2000).

Existem redes com diferentes fluxos de informações, sendo *feedforward* quando a informação passa da camada de entrada até chegar na camada de saída sem fazer laços, ou sendo recorrente quando a saída de um nó em uma camada mais profunda serve como entrada em um neurônio em uma camada mais rasa (BRAGA et al., 2000).

Em redes multicamadas, podem ser observados diferentes tipos de conexões, uma rede é completamente conectada quando os nós de uma camada estão conectados com todos os nós das camadas anteriores e posteriores, parcialmente conectada quando os neurônios estão conectados com alguns neurônios das camadas anteriores e posteriores e por fim, quando um neurônio é conectado a outros neurônios em uma localização bem definida, é chamada de localmente conectada (FACELI et al., 2011).

2.2.10 Aprendizado das Redes Neurais Artificiais

Segundo Braga et al. (2000) a primeira fase de uma RNA para sua atuação é

a aprendizagem, é nela que se apresenta o domínio do problema e revela-se padrões de informações utilizados para criar uma representação própria da entrada. Haykin (2001) diz que uma RNA se torna mais poderosa à medida que aprende sobre o domínio de seu problema, este processo é realizado através de um método iterativo onde ocorre o ajuste dos pesos das entradas dos neurônios. Faceli et al. (2011) dizem que para realizar o ajuste de pesos são utilizados de algoritmos que estabelecem quando e como devem-se alterar pesos para aumentar a capacidade de previsão da rede.

Existe um conjunto de técnicas para realizar o aprendizado da rede, mas não existe uma ferramenta universal de aprendizado, pois cada método tem suas diferenças e, como consequências, apresentando vantagens e desvantagens entre si. existem quatro regras básicas de aprendizagem: aprendizagem por correção de erro (1), hebbiana (2), competitiva (3) e aprendizagem de Boltzmann (4) (HAYKIN, 2001).

Segundo Faceli et al. (2011), na aprendizagem por correção de erro (1) se reduz os erros cometidos da rede ajustando os pesos das RNAs, sendo um tipo de aprendizado supervisionado. Na aprendizagem hebbiana (2) segue-se a regra de Hebb onde neurônios que emitem sinais ao mesmo tempo tendem a ter uma conexão sináptica mais forte do que os neurônios que emitem sinais assincronamente, sendo um tipo de aprendizado não supervisionado. Na aprendizagem competitiva (3) os neurônios competem para serem ajustados, os vencedores que tem seus pesos ajustados, pois respondem melhor ao problema apresentado, sendo também um tipo de aprendizado não supervisionado e, por fim, os termodinâmicos de Boltzmann (4) no qual contém princípios termodinâmicos.

2.2.11 Algoritmo de Treinamento *Backpropagation*

Aumentar a quantidade de camadas e neurônios de uma RNA trouxe consigo o problema de descobrir como corrigir cada conexão quando se erra a resposta à uma instância em seu treinamento. Este dilema foi abordado como o problema da atribuição de crédito por Coppin (2010), Faceli et al. (2011).

O problema da atribuição de crédito foi resolvido pelo *Backpropagation* que faz a análise do conjunto de entrada com o da saída desejada apresentados à rede a cada iteração de treinamento. Esse processo apenas termina quando um critério de parada é atingido no qual este critério pode ser uma condição arbitrária definida pelo projetista da rede (FACELI et al., 2011; COPPIN, 2010).

O funcionamento *Backpropagation* ocorre em duas fases: a fase *forward* que alimenta a rede com dados para que ela chegue a uma saída e a fase *backward* onde o fluxo de informação passa da camada de saída para a de entrada usando o erro encontrado para corrigir os pesos das conexões entre os neurônios (BRAGA et al., 2000).

Um requisito para se utilizar do *Backpropagation* é o uso de funções de ativações pelos nós que sejam diferenciáveis e de preferência não decrescente, ou seja, funções de ativações não lineares (BRAGA et al., 2000).

2.2.12 Perceptron Multicamadas

Segundo Coppin (2010) o Perceptron é uma ótima rede para estudar o neurônio artificial, porém sua limitação para problemas linearmente separáveis é um grave problema para sua aplicação no mundo real, esta seção busca descrever o Perceptron Multicamadas.

O Perceptron Multicamadas foi resultado do esforço de fazer as RNAs resolverem problemas não linearmente separáveis. Para isso ela é *feedforward* podendo ter uma ou mais camadas com inúmeros neurônios completamente conectados em que é preciso uma função de ativação diferenciável para que seja possível a aplicação do algoritmo de treinamento *backpropagation* (FACELI et al., 2011).

A quantidade de nós da rede é definido empiricamente, sendo utilizado como parâmetros os números de entradas e saídas da rede, número de exemplos de treinamento, quantidade de ruídos pertencentes ao exemplo a complexidade do problema (BRAGA et al., 2000).

Cada neurônio desta rede realiza uma tarefa específica, e a medida que a rede se aprofunda esta função se torna mais complexa e necessita de mais processamento. Pode-se interpretar a função dos neurônios da seguinte maneira: na primeira camada oculta cada neurônio delimita um hiperplano, os neurônios da camada seguinte combinam uma região convexa com as delimitações da primeira camada e os neurônios da camada de saída formam subconjuntos destas regiões convexas criadas pela penúltima camada (FACELI et al., 2011).

Um grande problema da MLP é seu treinamento, pois a única camada que tem uma noção completa do erro é a de saída, na última camada oculta é possível obter a estimativa do erro para cada neurônio e na penúltima camada oculta apenas a estimativa do erro repassado da última camada, e o problema se agrava se a rede possui mais camadas ocultas (BRAGA et al., 2000).

Na camada de saída cada neurônio representa uma classe, para fazer a comparação do erro é feito a comparação vetorial dos resultados de saída real da rede com a saída desejada para aquele conjunto de entrada (FACELI et al., 2011). Braga et al. (2000) relatam que garantir a função de treino da rede multicamada não significa na aplicabilidade real desta pois o conjunto de treinamento pode resultar num mínimo local da função real.

2.2.13 Redes Neurais Artificiais Recorrentes

As redes neurais recorrentes são uma extensão das RNAs para resolver problemas relacionados a sequências com incorporações arbitrárias em sua arquitetura de conexões recorrentes pois quase sempre as atividades humanas estão ligadas à uma cadeia de acontecimentos por exemplo no processamento de linguagem natural, memórias ou atividades motoras que são ações sequenciais (ROSA, 2011).

É possível abstrair a ideia de continuidade de abordagens recorrentes em redes MLP não recorrentes quando se adiciona entradas que representam o efeito temporal em seu processamento, ou seja, pode-se utilizar o tempo como um dos parâmetros de entrada da rede, porém este método pode aumentar a dimensionalidade do problema o que ocasiona na maldição da dimensionalidade e não garante matematicamente que a saída atual da rede dependa dos resultados anteriores (ROSA, 2011; MANASWI, 2018).

A arquitetura da rede recorrente garante matematicamente que o resultado dependa dos últimos resultados computados pela rede como visto na Equação 2.7 que explica a relação das conexões recorrentes da Figura (4) (MANASWI, 2018).

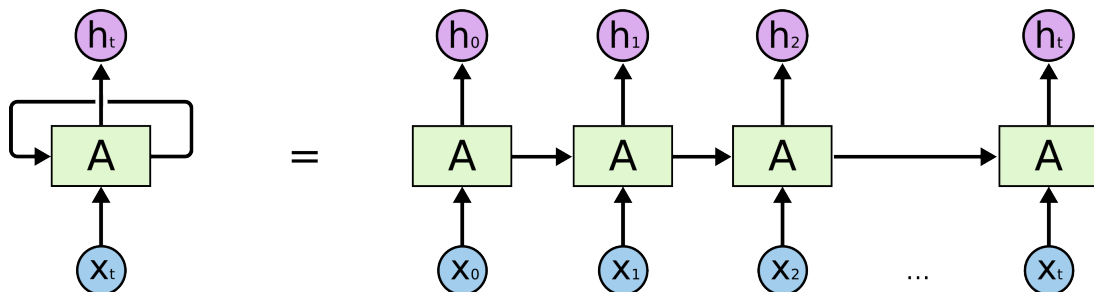


Figura 4 – Rede neural recorrente.

Fonte: (COLAH, 2011)

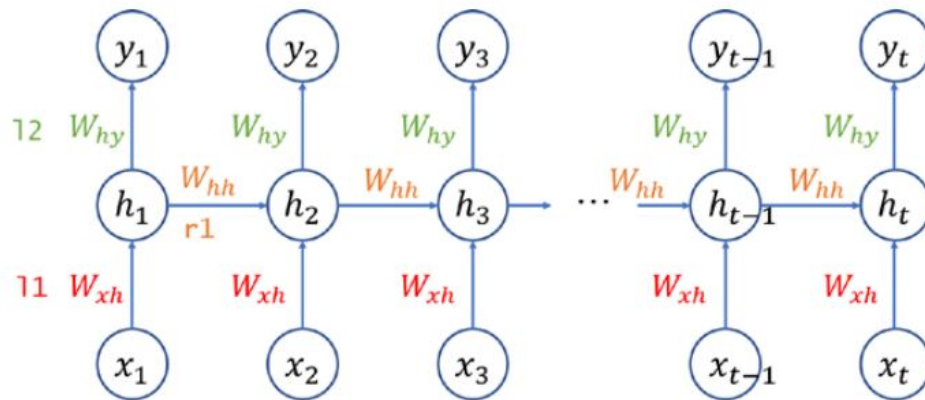


Figura 5 – Rede neural recorrente com os pesos.

Fonte: (MANASWI, 2018)

$$\begin{aligned} h_t &= \tanh(l1(x_t) + r1(h_{t-1})) \\ y_t &= l2(h_t) \end{aligned} \quad (2.7)$$

Onde:

h_t = passo temporal no tempo t ;

$l1$ = função de ativação aplicada na entrada da rede neural recorrente;

x_t = entrada da rede neural recorrente no tempo t ;

$r1$ = função de ativação aplicada na saída obtida do processamento no tempo $t - 1$;

h_{t-1} = saída da rede neural recorrente obtida no tempo $t - 1$;

y_t = saída da rede neural recorrente no tempo t ;

$l2$ = função de ativação aplicado no potencial de ativação da rede neural recorrente.

Para treinar uma rede neural recorrente utiliza-se o algoritmo de retropropagação através do tempo, no qual a rede neural recorrente é transformada em uma rede neural não recorrente desdobrando os passos temporais em novas camadas (Figura (5)) ação que faz ser possível aplicar o algoritmo backpropagation normalmente (HAYKIN, 2001; MEDSKER; JAIN, 2001).

Desdobrar uma rede neural recorrente de muitos passos temporais em uma rede neural não recorrente cria uma rede profunda que faz agravar o problema da dissipação do gradiente no qual as memórias de longo prazo são ocultadas pelas memórias de curto prazo (MANASWI, 2018; HOCHREITER; SCHMIDHUBER, 1997; NELSON, 2017).

2.2.14 A RNA Long Short Term Memory

Para solucionar o problema da dissipação do gradiente e dar a devida importância às memórias de longo prazo Hochreiter e Schmidhuber (1997) apresentaram a rede LSTM, uma RNA recorrente que contém um estado celular regulada por uma estrutura de portões que segundo Manaswi (2018) são responsáveis por decidir se as informações de longo prazo devem ser esquecidas ou se as novas informações do tempo atual da rede deve ser guardada para serem usadas posteriormente.

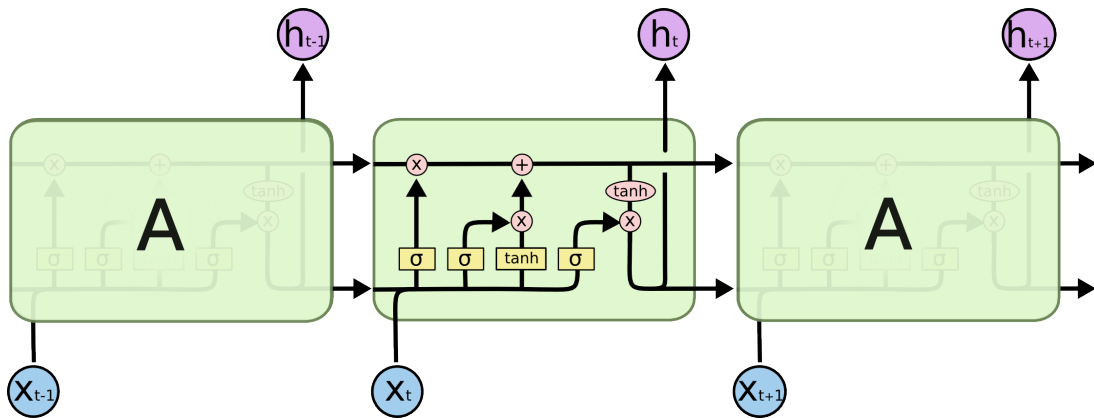


Figura 6 – Rede Long Short Term Memory.

Fonte: (COLAH, 2011)

Na Figura 6 contém um exemplo de um nó de processamento LSTM. Nele, se vê o estado celular atravessando todos os tempos no canto superior do nó, vê-se também o passo atual da rede que é a entrada do novo parâmetro da rede de X_t com o último processamento da rede h_{t-1} passando por portões que tem a responsabilidade de atualizar o estado celular e processar a saída da rede.

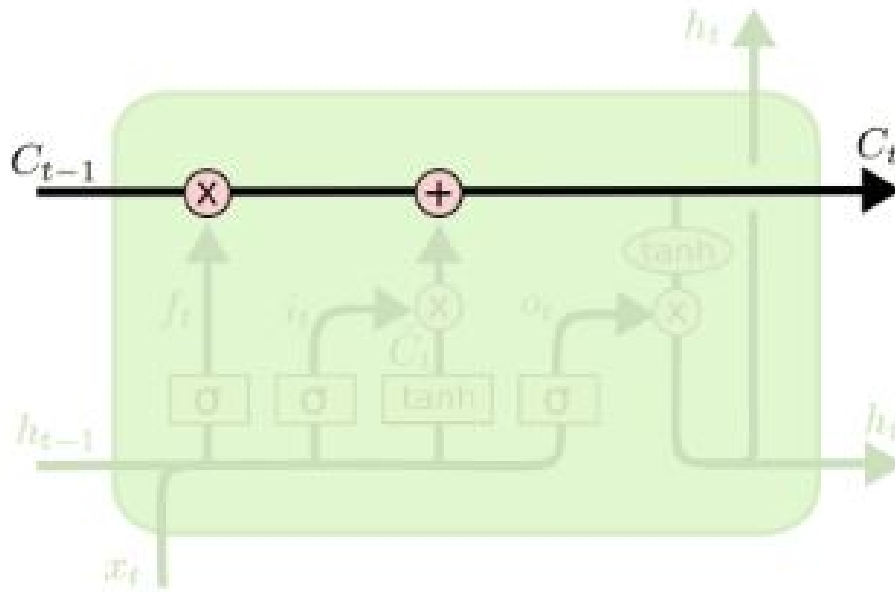


Figura 7 – Célula de estado LSTM.

Fonte: (COLAH, 2011)

A célula de estado (Figura (7)) é um dos principais componentes para o funcionamento da rede LSTM, nele é adicionado ou removido informações da sequência por operações de multiplicação e adição (Equação(2.8)) regulados pelos portões de esquecimento f_t e de entrada $i_t \times \tilde{C}_t$ amenizando o problema da dissipação do gradiente (HOCHREITER; SCHMIDHUBER, 1997; COLAH, 2011)].

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (2.8)$$

No que:

C_t = célula de estado no tempo t ;

f_t = resultado obtido do portão do esquecimento t ;

C_{t-1} = célula de estado no tempo $t - 1$;

i_t = função Sigmoide Logística aplicada na entrada da LSTM no tempo t em conjunto com o estado da LSTM no tempo $t - 1$.;

\tilde{C}_t = função tanh aplicada na entrada da LSTM no tempo t em conjunto com o estado da LSTM no tempo $t - 1$.

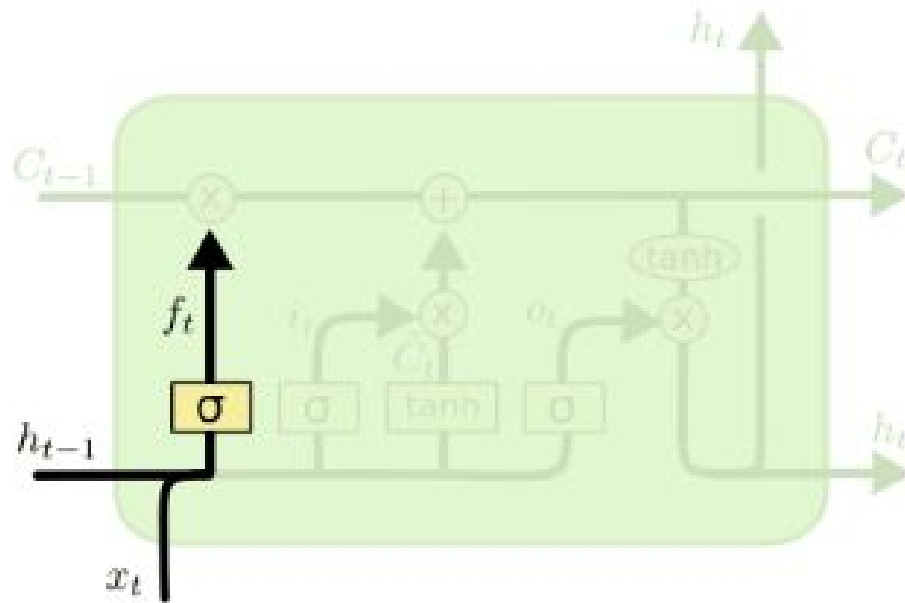


Figura 8 – Portão do esquecimento LSTM.

Fonte: (COLAH, 2011)

O portão do esquecimento (Figura 8) tem função de decidir o que será esquecido do estado celular, para isso usa uma função sigmoide logística σ que tem como parâmetro a entrada atual da rede X_t e sua memória h_{t-1} , multiplicando a sua saída f_t no estado celular C_{t-1} (NELSON, 2017).

$$f_t = \sigma(Wf[h_{t-1}, X_t]) \quad (2.9)$$

Onde:

- σ = função de ativação Sigmoide Logística;
- X_t = entrada da rede LSTM no tempo t ;
- h_{t-1} = saída da rede LSTM no tempo $t - 1$;
- W = peso da conexão do portão do esquecimento.

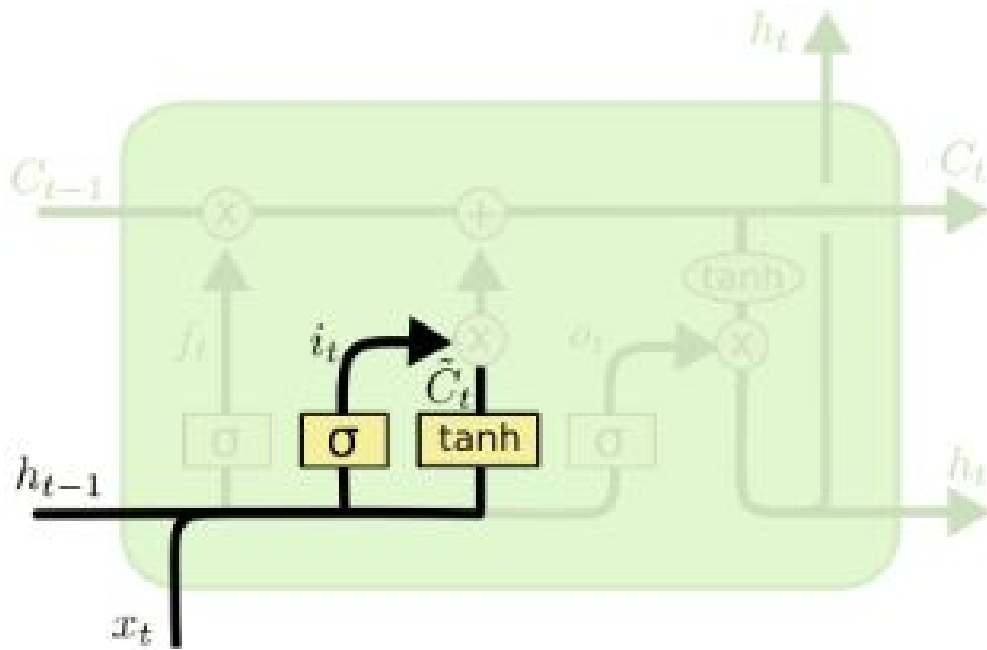


Figura 9 – Portão de entrada LSTM.

Fonte: (COLAH, 2011)

O portão de entrada (Figura 9) adiciona informações relevantes na célula de estado C_{t-1} pela multiplicação do resultado de uma função sigmoide logística σi_t (Equação(2.10)) por \tilde{C}_t que é a saída da função tangente hiperbólica \tanh (Equação(2.11)). Ambas as funções σ e \tanh possuem h_{t-1} e x_t como parâmetros (HOCHREITER; SCHMIDHUBER, 1997).

$$i_t = \sigma(Wf[h_{t-1}, x_t]) \quad (2.10)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t]) \quad (2.11)$$

No que:

i_t = Sigmoide logística que processa o resultado da LSTM no tempo $t - 1$ com a entrada no tempo t ;

\tilde{C}_t = resultado da \tanh que recebe como parâmetro resultado da LSTM no tempo $t - 1$ com a entrada no tempo t ;

W_c = conjunto de pesos que conecta a entrada do tempo t e o resultado da LSTM do tempo $t - 1$ antes de ser usado processada pela \tanh ;

W = conjunto de pesos que conecta a entrada do tempo t e o resultado da LSTM do tempo $t - 1$ antes de ser usado processada pela Sigmoide logística.

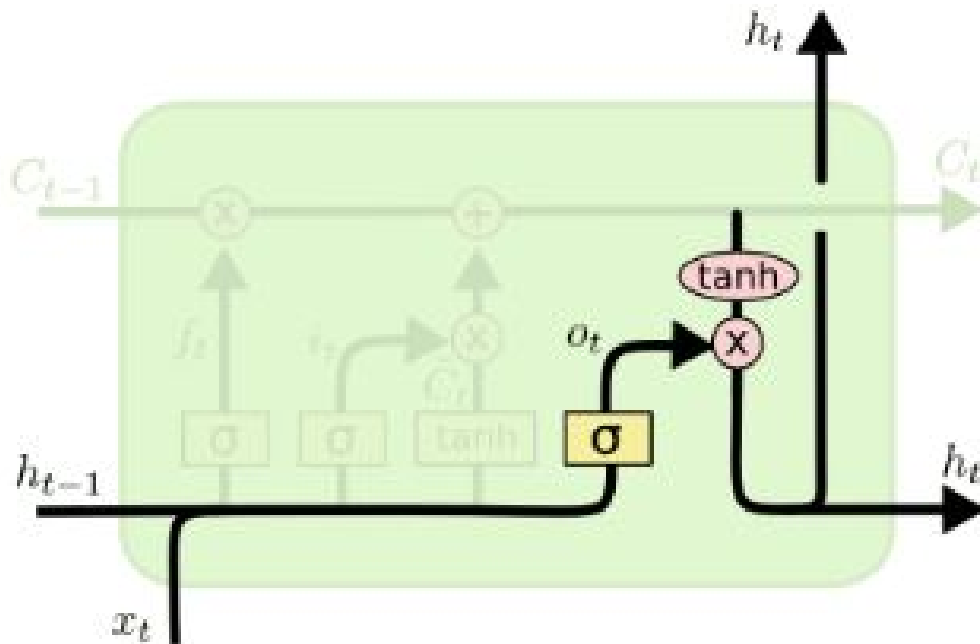


Figura 10 – Saída da LSTM.

Fonte: (COLAH, 2011)

Para determinar a saída da rede h_t (Figura 10), aplica-se uma função sigmoide logística σ à entrada X_t e à última saída da rede h_{t-1} (Equação(2.12)), e multiplica seu resultado O_t pela função tangente hiperbólica \tanh que recebe o estado celular C_t como parâmetro (Equação(2.13)) (BEIL et al., 2015).

$$O_t = \sigma(W_o[h_{t-1}, x_t]) \quad (2.12)$$

$$\tilde{C}_t = \tanh(C_t) \quad (2.13)$$

$$h_t = O_t \times \tilde{C}_t \quad (2.14)$$

Sendo que:

O_t = Sigmoide logística que processa o resultado da LSTM no tempo $t - 1$ com a entrada no tempo t ;

W_o = conjunto de pesos que conecta a entrada do tempo t e o resultado da LSTM do tempo $t - 1$;

\tilde{C}_t = resultado da \tanh que recebe como parâmetro o estado celular da LSTM;

h_t = saída da LSTM no tempo t .

3 MATERIAIS E MÉTODOS

O capítulo de materiais e métodos tem como objetivo apresentar os materiais e apontar as metodologias utilizadas neste trabalho. A (Figura 11) mostra os passos necessários para realizar o experimento.

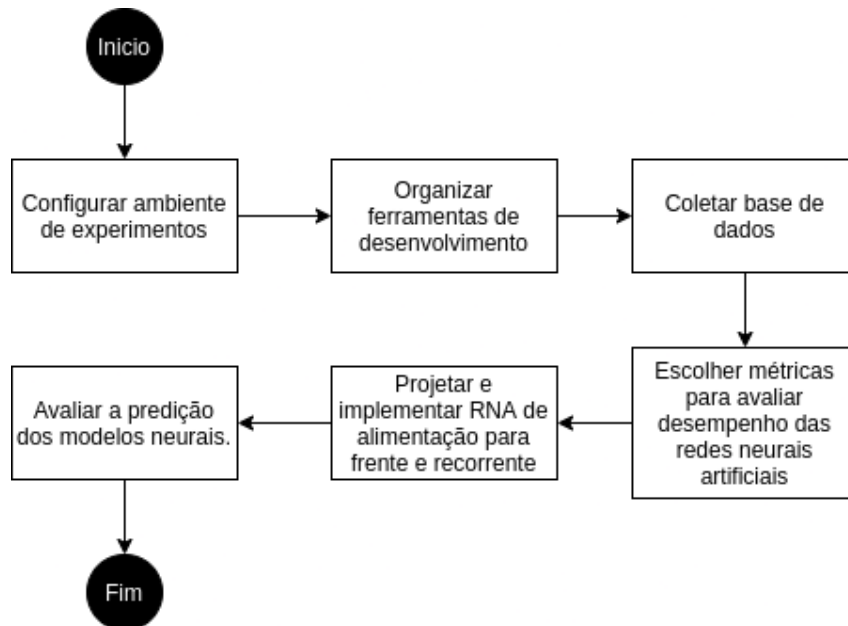


Figura 11 – Passos necessários para realizar o experimento.

Fonte: Autoria própria

3.1 CONFIGURAR AMBIENTE DE EXPERIMENTOS

Foi necessária a utilização de um servidor da UTFPR com uma máquina virtual instalada com a distribuição Linux Ubuntu versão (16.04) como seu sistema operacional para viabilizar o desenvolvimento do experimento desta monografia. O Ubuntu é um sistema

operacional de código aberto desenvolvido pela Canonical e muito usado em computadores e servidores.

Um experimento de aprendizado de máquina normalmente requer muitos recursos computacionais. Para executar os experimentos, foi alocado na máquina virtual 16 GB de memória RAM, 357 GB de armazenamento e 4 núcleos de processamento com 2.7 GHz.

O acesso ao servidor foi realizado pelo AnyDesk que é um aplicativo especializado em fornecer conectividade entre dois computadores remotos de modo que após instalar o AnyDesk no servidor e guardar o número de identificação e senha de acesso foi possível utilizar os recursos disponíveis do servidor por um ponto de acesso remoto.

3.2 ORGANIZAR FERRAMENTAS DE DESENVOLVIMENTO

O Python foi a linguagem de programação utilizada para tratar os dados coletados e desenvolver as redes neurais utilizadas neste trabalho. Python é uma linguagem de programação livre e multiplataforma na qual existe uma vasta gama de ferramentas voltadas para o aprendizado de máquina, por exemplo a *Application Programming Interface*(API) Keras.

A API Keras é designada para permitir o rápido desenvolvimento de RNAs, sendo idealizada para ser uma interface amigável ao usuário e atua utilizando *backends* tais como Tensorflow, Keano e *Microsoft Cognitive Toolkit* para lidar com a parte dos cálculos matemáticos.

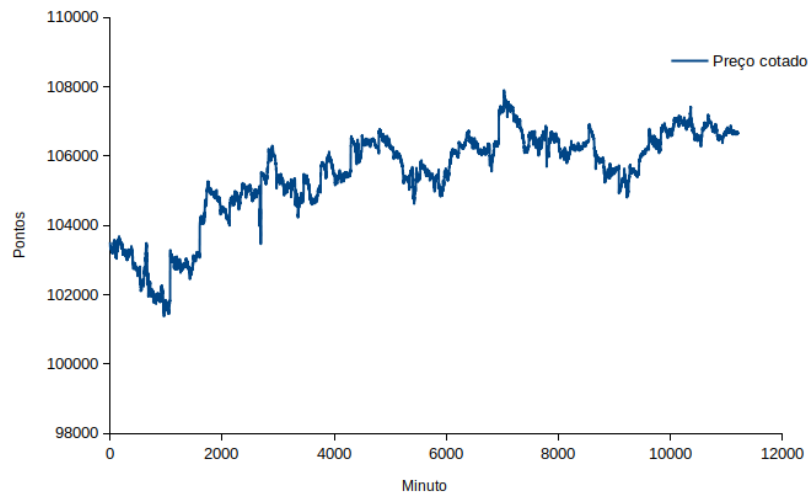
O Tensorflow é o *backend* usado para dar suporte ao Keras. A sua função é fazer computação numérica em grafos de modo que cada nó recebe um tensor e faz uma operação matemática, sendo o responsável pela parte de cálculos matemáticos complexos enquanto o Keras é o *frontend* que facilita a sua manipulação. A Keras e o Tensorflow são acessados nesse trabalho no ambiente de programação da linguagem Python para a criação dos modelos de RNAs.

3.3 COLETAR BASE DE DADOS

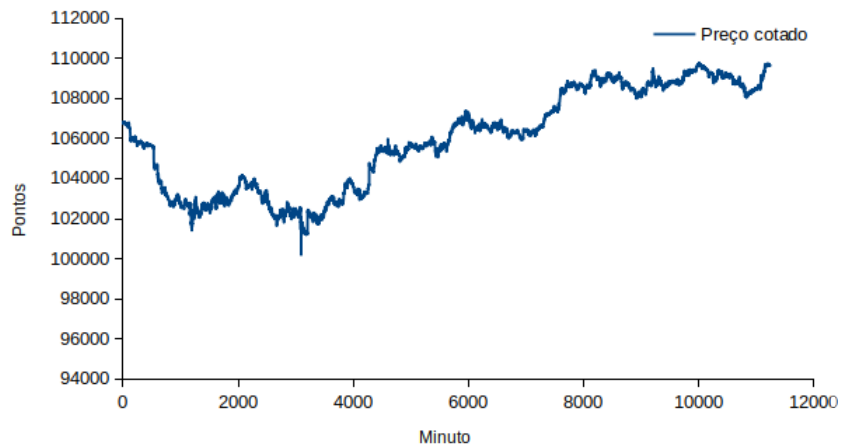
Segundo Fayyad et al. (2001), o processo de construção da base de dados é uma das etapas mais importantes de um projeto de aprendizado de máquina. Para a coleta do histórico de cotações do mini índice Bovespa foi utilizado o software *Metatrader 5*, um sistema utilizado para empreender negociações *online* que possui ferramentas gráficas para a prática de análise técnica no preço além de permitir o desenvolvimento de robôs de negociação automatizada. Por fim, a coleta de dados foi salva em um arquivo do tipo *Comma Separated Value*(CSV) no qual cada linha do arquivo representa uma instância do conjunto de dados.

Para este trabalho, foi desenvolvida uma base de dados com informações do preço cotado do contrato futuro do mini Índice Bovespa entre 30/10/2017 até 07/12/2017 no qual é usado para encontrar as configurações dos modelos neurais que melhor se encaixam ao conjunto de dados sendo mantido a mesma sazonalidade da época do teste dos preditores porém numa diferença de dois anos para evitar que ocorresse um sobreajuste das melhores configurações. Para o treinamento e validação dos melhores modelo foi selecionado o preço cotado entre os dias 01/09/2019 a 31/10/2019 e para o teste dos modelos foi escolhido o preço negociado entre 01/11/2019 a 30/11/2019 sendo as datas escolhidas o mais próximo do período de teste para fazer com que as redes neurais aprendessem como o mercado esteve se comportando para poder realizar predições mais acuradas.

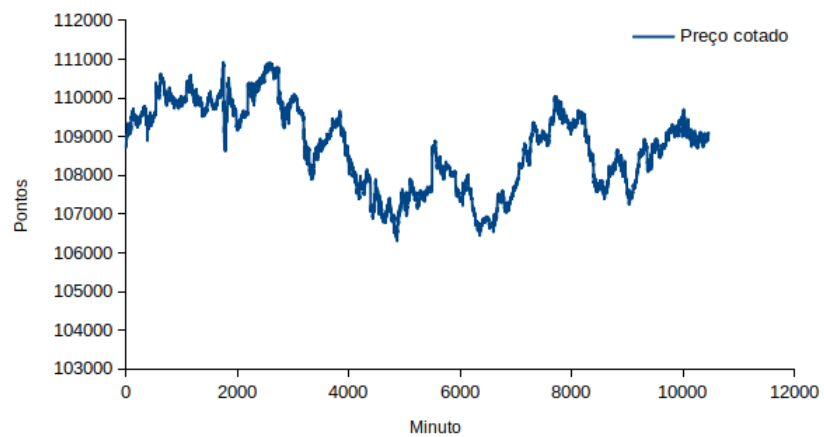
Os gráficos de linha da Imagem 12 servem para mostrar como estava a tendência no mercado durante o período de treinamento (Imagem 12a), validação (Imagem 12b) e teste (Imagem 12c) dos modelos de alimentação recorrente e adiante, se referindo a como os preços do contrato futuro do mini Índice Bovespa foram negociados nestes períodos.



(a) Treinamento



(b) Validação



(c) Teste

Figura 12 – Cotação do contrato futuro do mini índice Bovespa nos conjuntos de: a) treinamento, b) validação, c) teste.

Fonte: Autoria própria

O tipo de gráfico mais usado pelos investidores em bolsas de valores é o gráfico de vela no qual cada vela se refere à um período constante de tempo, sendo possível guardar informações como o preço máximo, mínimo, de abertura e de fechamento que foi cotado durante o período de tempo relacionado à vela.

A base de dados utilizada neste trabalho toma como base um gráfico de vela de período definido de um minuto, e por isso, conta com atributos que informam o preço de abertura, máxima, mínima e fechamento além do dia da semana, hora e minuto referente ao preço cotado no intervalo de minuto a minuto do início ao fim da negociação diária na bolsa de valores. O preço do contrato futuro do Mini Índice Bovespa é cotado em pontos no qual cada ponto equivale a 0,2R\$. Ao multiplicar a quantidade de pontos pelo valor do ponto é possível chegar ao valor monetário contrato.

Foi decidido empiricamente que as amostras que são apresentadas para os modelos de RNA recorrente e não recorrente possuem como domínio uma matriz de 200 instâncias com os atributos de abertura, máxima, mínima, dia da semana, hora e minuto (Tabela 1) e como imagem um vetor com o fechamento do preço nos próximos 60 minutos marcados do fim do domínio da amostra (Tabela 2).

Tabela 1 – Exemplo de 10 instâncias do domínio do conjunto de treinamento

ABERTURA	MÁXIMA	MÍNIMA	DIA_SEMANA	HORA	MINUTO
79902	79912	79844	Seg	09	00
79886	79941	79826	Seg	09	01
79916	79996	79901	Seg	09	02
79936	79966	79866	Seg	09	03
79876	79896	79731	Seg	09	04
79761	79821	79751	Seg	09	05
79821	79821	79791	Seg	09	06
79806	79806	79731	Seg	09	07
79746	79796	79736	Seg	09	08
79786	79816	79781	Seg	09	09

Fonte: Autoria própria

Tabela 2 – Exemplo de 10 instâncias da imagem do conjunto de treinamento

FECHAMENTO
79.871
79.856
79.776
79.811
79.771
79.786
79.796
79.801
79.776
79.761

Fonte: Autoria própria

Os atributos da base de dados são quantitativos, então foi realizada a normalização dos dados para que fiquem entre 0 e 1 e assim possuïrem a mesma escala. No caso do dia da semana, que é um atributo ordinal, primeiro foi feita a transformação para atributo numérico e então normalização.

3.4 ESCOLHER MÉTRICAS PARA AVALIAR O DESEMPENHO DAS RNAS

Neste trabalho foi utilizado o *Mean Squared Error*(MSE), coeficiente de correlação de Pearson e o coeficiente de determinação R^2 . Na equação (3.1) e (3.3) Y_i e \hat{Y}_i significam a saída da rede com a saída desejada enquanto na equação (3.2) x_i significa uma instância de uma amostra enquanto y_i é uma instância de uma outra amostra.

O MSE é a média das somas dos erros ao quadrado cometidos pela rede (equação 3.1), o coeficiente de correlação de Pearson (equação 3.2) mede a correlação entre duas variáveis em que resulta em 1 quando são totalmente relacionadas, -1 quando existe uma relação inversa entre as duas variáveis ou em 0 quando não há relação entre as variáveis. O R^2 varia de 0 a 1, é uma métrica que indica o quanto o modelo regressivo consegue explicar as amostras, no qual 0 representa que o modelo não consegue explicar as amostras enquanto 1 consegue explicá-las

perfeitamente (equação 3.3).

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (3.1)$$

Sendo que:

n = número de observações;

Y_i = saída obtida pelo modelo avaliado após processar uma instância apresentada;

\hat{Y}_i = saída esperada da rede para o a instância apresentada.

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^n (x_i - \bar{x}_i)^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y}_i)^2}} \quad (3.2)$$

Onde:

n = número de observações;

x_i = uma instância da amostra x ;

\bar{x}_i = média da amostra x ;

y_i = uma instância da amostra y ;

\bar{y}_i = média da amostra y .

$$\begin{aligned} R^2 &= 1 - \frac{SQ_{res}}{SQ_{tot}} \\ SQ_{res} &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\ SQ_{tot} &= \sum_{i=1}^n (Y_i - \bar{Y}_i)^2 \end{aligned} \quad (3.3)$$

Em que:

SQ_{res} = soma dos quadrados dos resíduos;

SQ_{tot} = soma de quadrados totais;

n = número de observações;

\hat{Y}_i é o valor estimado (previsão) de Y_i ;

\bar{Y}_i = média das observações.

3.5 PROJETAR E IMPLEMENTAR RNA DE ALIMENTAÇÃO RECORRENTE E ADIANTE

Foi realizado a descida coordenada de hiperparâmetros que consiste em um teste de dois passos nos experimentos subsequentes para encontrar os melhores modelos preditivos. O primeiro passo consiste em achar o número de camada que obtém o melhor desempenho e com

ele iniciar o segundo passo em que é almejado a otimização da taxa de aprendizado.

Em ambos os passos devido ao grande esforço de tempo que seria necessário para testar topologias com diferentes números de neurônios e funções de ativação por camada foi mantido constante a quantidade de neurônios das camadas ocultas em 64 e sigmoid logística como função de ativação nas camadas ocultas assim como foi mantido 60 neurônios e função linear como função de ativação na camada de saída.

O processo de decisão da melhor topologia a dar seguimento aos testes foi definido pelo modelo que obteve as melhores métricas. Tanto no primeiro como no segundo passo, devido à grande quantidade de dados apresentados às redes e à necessidade de analisar apenas qual topologia melhor se adaptava ao conjunto de dados foi decidido utilizar três épocas para treinar os modelos. No primeiro passo, seguindo o algoritmo descida coordenada de hiperparâmetros foram testados seis modelos de RNA incrementando-se uma camada oculta ao modelo em cada iteração.

A taxa de aprendizado padrão do algoritmo de treinamento do Keras é 0.001. Por isso, a otimização da taxa de aprendizado no segundo passo gravitou em torno deste número com o objetivo de minimizar o MSE. Após mais de seis testes orientados pelo declínio do MSE, foi utilizado o mesmo teste de métricas do primeiro passo para definir a rede recorrente e não recorrente que melhor abstrai o problema de predição.

3.6 AVALIAR A PREDIÇÃO DOS MODELOS NEURAIIS

O melhor modelo recorrente e *feedforward* foram usados na predição da cotação do mini Índice Bovespa na amostra de dados do mês de novembro de 2019. Ao aplicar o melhor modelo de alimentação recorrente e para frente para prever cada minuto negociado no mês de novembro de 2019 foi realizada 10261 predições e com elas foram realizadas a análise de correlação de Pearson com os valores de cotações reais. Para analisar as predições foi usado apenas a predição do próximo minuto e por isso a camada de saída de ambos os modelos neurais foi alterada para que tenha apenas um neurônio.

Foi escolhido empiricamente dois preditores para serem usados como grupo de controle: um que prediz que o próximo valor seja a média do preço fechado dos últimos 200 minutos, e um segundo preditor que sempre prediz que não haverá alteração de preço no próximo minuto (preditor *baseline*).

4 RESULTADOS OBTIDOS E DISCUSSÃO

Tendo definido os materiais e métodos no capítulo anterior, foram iniciados os preparativos para que se tenha uma rede de alimentação recorrente e outra de alimentação para frente bem treinadas para depois realizar a coleta e análise de suas predições para as cotações apresentadas às RNAs referente ao mês de Novembro em 2019.

4.1 ELABORAÇÃO E CONCEPÇÃO DO MODELO NÃO RECURRENTE

A descida coordenada de hiper-parâmetros foi iniciada no modelo de RNA de alimentação para frente e no primeiro passo do teste a rede de duas camadas foi a escolhida para ser testada com diferentes taxas de aprendizado, pois apenas não obteve a melhor métrica no MSE do conjunto de validação (Tabela 3).

Tabela 3 – Resultado do teste de topologias de redes neurais artificiais de alimentação para frente de 30/10/2017 à 07/12/2017

Camadas	Mse	Validação Mse	Pearson	Validação Pearson	R ²	Validação R ²
1	0,00776	0,00954	0,8890	0,8927	0,7904	0,7969
2	0,00331	0,00527	0,9407	0,9438	0,8849	0,8909
3	0,00354	0,00405	0,9353	0,9395	0,8748	0,8827
4	0,00369	0,00569	0,9328	0,9366	0,8702	0,8773
5	0,00392	0,00389	0,9235	0,9291	0,8530	0,8633
6	0,00415	0,00549	0,9164	0,9224	0,8399	0,8508

Fonte: Autoria própria

A topologia de duas camadas é a que melhor abstrai o problema ao obter o menor

valor para MSE no conjunto de treino e o maior valor do coeficiente de correlação de Pearson e determinação R^2 nos conjuntos de treino e validação, então, foi dado início ao segundo passo do teste que são os experimentos com as taxas de aprendizados de 0.01, 0.005, 0.001, 0.0005, 0.00025, 0.0001 na topologia escolhida de duas camadas e foi escolhido o uso do decaimento exponencial nos valores da taxa de aprendizado para uma otimização das conexões neurais menos caótica. Ao fim do segundo passo, foi selecionada a taxa de aprendizado 0.0005 (Tabela 4) que obteve os maiores valores no coeficiente de correlação de Pearson e determinação R^2 no conjunto de treino e validação.

Tabela 4 – Resultado do teste de taxas de aprendizado de redes neurais artificiais de alimentação para frente de 30/10/2017 à 07/12/2017

	Mse	Validação Mse	Pearson	Validação Pearson	R^2	Validação R^2
0,01	0,00328	0,00429	0,9366	0,9408	0,8772	0,8852
0,005	0,00333	0,00373	0,9394	0,9434	0,8826	0,8901
0,001	0,00331	0,00527	0,9407	0,9438	0,8849	0,8909
0,0005	0,00325	0,00485	0,9420	0,9453	0,8875	0,8936
0,00025	0,00322	0,00478	0,9395	0,9432	0,8827	0,8897
0,0001	0,00333	0,00354	0,9414	0,9450	0,8863	0,8930

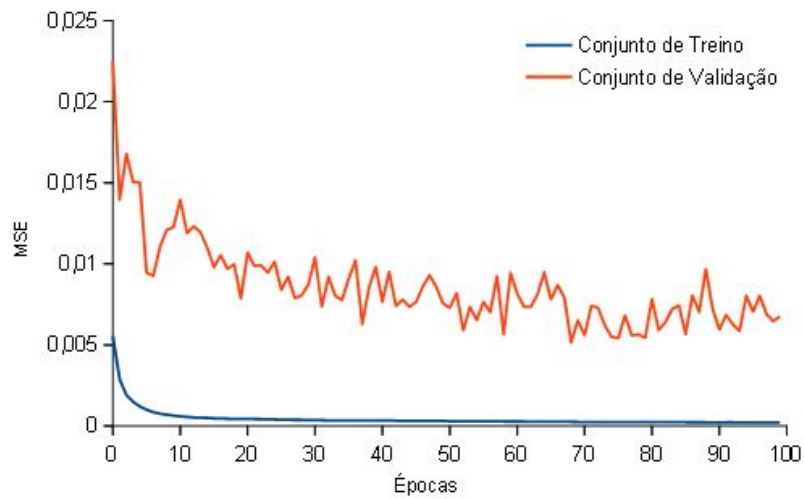
Fonte: Autoria própria

Com os testes de diferentes topologias e taxas de aprendizado terminado, foram decididos os parâmetros para serem utilizados na RNA não recorrente. Logo, a MLP possui duas camadas com a função de ativação sigmoid logística com uma taxa de aprendizado de 0.0005 atuando com uma quantidade de 62 neurônios na primeira e segunda camada além de 60 neurônios na camada de saída com uma função de ativação linear.

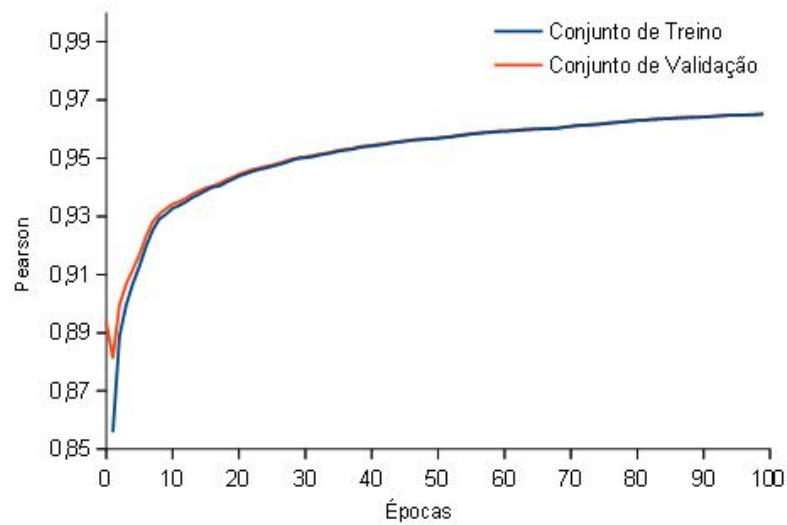
4.2 OTIMIZAÇÃO DO MODELO DE ALIMENTAÇÃO PARA FRENTE

Na otimização da rede de alimentação para frente é realizado um treinamento do melhor modelo neural de alimentação para frente obtido na Seção 4.1 no mês de setembro de 2019 e validando com o mês de outubro de 2019. Empiricamente foi definido um limite de

100 épocas para o treinamento, do histórico deste treinamento foi possível gerar gráficos sobre a variação do MSE, Correlação de Pearson e R^2 obtidos dos conjuntos de treinamento e validação do modelo durante o seu treinamento (Figura 13).



(a) Mean Square Error (MSE)



(b) Coeficiente de Pearson (R)

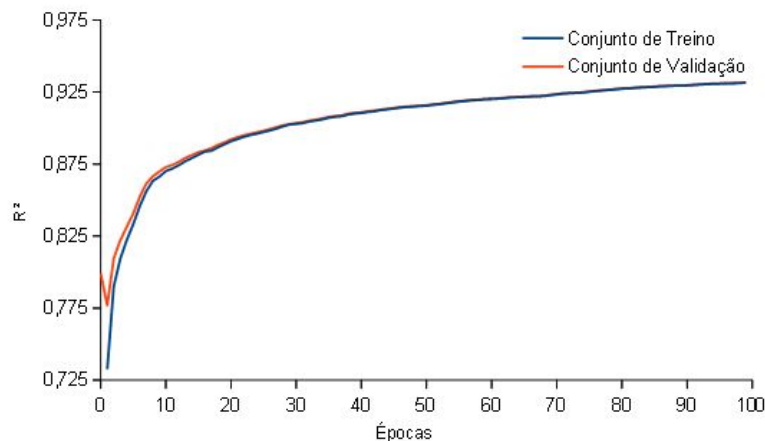
(c) Coeficiente de Determinação (R^2)

Figura 13 – Mean Square Error (MSE), Coeficiente de Pearson (R) e Coeficiente de Determinação (R^2) do período de treinamento e validação da rede de alimentação para frente.

Fonte: Autoria própria

Ao decorrer das épocas de seu treinamento, o MSE faz uma curva de redução até finalizar na centésima época com o valor de 0,00021 para o conjunto de treinamento e 0,0067 para o conjunto de validação. No quesito correlação de Pearson e R^2 , houve uma curva de crescimento até alcançar o 0,965 para a correlação de Pearson e 0,931 para o R^2 tanto no conjunto de treino como no de validação.

Os dados obtidos do treinamento indicam que houve boa abstração do problema pelo modelo neural de alimentação para frente, pois há a obtenção de valores muito próximos de zero para o MSE em um gráfico que esboça a redução desta métrica a cada nova época. Outro bom sinal foi o aumento da correlação de Pearson e também R^2 a um valor muito próximo de 1 que como descrito na Seção 3.4 1 é o maior valor que se pode esperar obter tanto para a correlação de Pearson como para o R^2 ao que indica um bom aprendizado do modelo.

4.3 ELABORAÇÃO E CONCEPÇÃO DO MODELO RECORRENTE

Com os parâmetros de RNA MLP, foi iniciado o processo de modelagem da RNA LSTM. Foi realizado o primeiro passo do teste e a rede de duas camadas LSTM (Tabela 5) obteve as melhores métricas no coeficiente de Pearson e no coeficiente de determinação R^2 nos conjuntos de treinamento e validação e no MSE do conjunto de validação.

Tabela 5 – Resultado do teste de topologias de redes neurais artificiais de alimentação recorrente de 30/10/2017 à 07/12/2017

Camadas	Mse	Validação Mse	Pearson	Validação Pearson	R ²	Validação R ²
1	0,00160	0,00302	0,9753	0,9758	0,9513	0,9523
2	0,00161	0,00255	0,9757	0,9764	0,9520	0,9534
3	0,00171	0,00290	0,9729	0,9739	0,9465	0,9485
4	0,00179	0,00384	0,9710	0,9719	0,9429	0,9446
5	0,00180	0,00321	0,9689	0,9703	0,9388	0,9415
6	0,00187	0,00300	0,9672	0,9688	0,9355	0,9387

Fonte: Autoria própria

A topologia recorrente que melhor abstrai o problema é uma rede com duas camadas ocultas por ter obtido o menor valor para o MSE no conjunto de validação e os maiores valores para o coeficiente de correlação de Pearson e determinação R² tanto no conjunto de treinamento como de validação. Com a melhor topologia escolhida foi iniciado o segundo passo do teste que são os experimentos com as taxas de aprendizados de 0.01, 0.005, 0.0025, 0.001, 0.00075, 0.0005 e 0.0001.

Ao fim do segundo passo foi selecionada a taxa de aprendizado 0.001 (Tabela 6) que obteve o menor valor para o MSE e os maiores valores para o coeficiente de correlação de Pearson e determinação R² no conjunto de treino e validação.

Tabela 6 – Resultado do teste de taxas de aprendizado de redes neurais artificiais de alimentação recorrente de 30/10/2017 à 07/12/2017

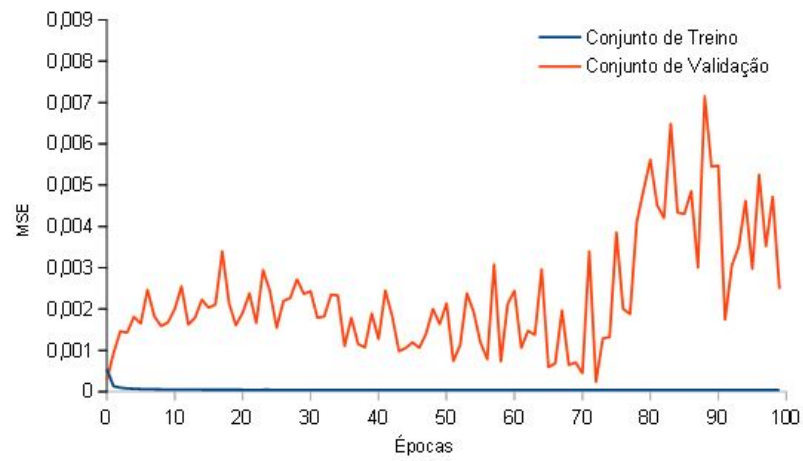
	Mse	Validação Mse	Pearson	Validação Pearson	R ²	Validação R ²
0,01	0,00167	0,00287	0,9743	0,9751	0,9494	0,9509
0,005	0,00167	0,00263	0,9742	0,9750	0,9491	0,9507
0,0025	0,00168	0,00264	0,9732	0,9742	0,9471	0,9491
0,001	0,00161	0,00255	0,9757	0,9764	0,9520	0,9534
0,00075	0,00166	0,00297	0,9737	0,9746	0,9482	0,9499
0,0005	0,00168	0,00258	0,9741	0,9749	0,9489	0,9505
0,0001	0,00166	0,00329	0,9744	0,9750	0,9496	0,9506

Fonte: Autoria própria

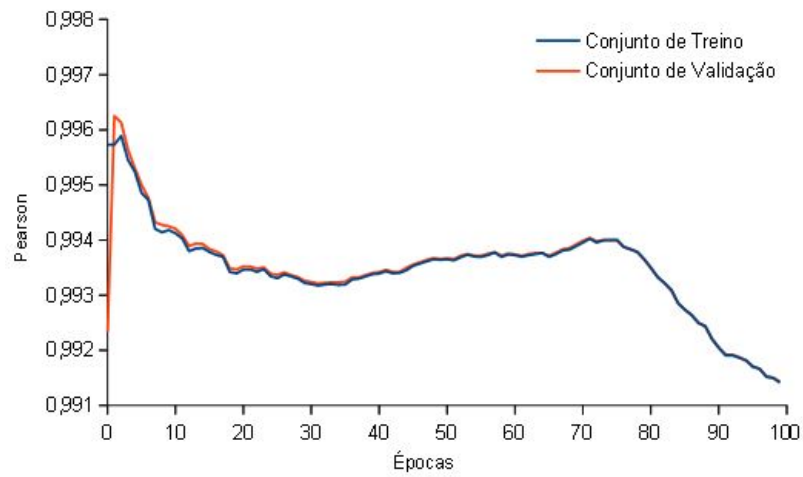
Ao fim do algoritmo de descida coordenada de hiper-parâmetros para encontrar o melhor modelo neural de alimentação recorrente foi resolvido que a RNA recorrente que deve possuir 2 camadas LSTM com a função de ativação Sigmoide Logística com 62 neurônios cada e uma camada densa (completamente conectada) como saída com 60 neurônios e função de ativação linear, todas com 0.001 como taxa de aprendizado.

4.4 OTIMIZAÇÃO DO MODELO RECORRENTE

Repetindo o processo de otimização da rede com alimentação para frente para a rede recorrente, obteve-se os gráficos das variações do MSE, correlação de Pearson e coeficiente de determinação.



(a) Mean Square Error (MSE)



(b) Coeficiente de Pearson (R)

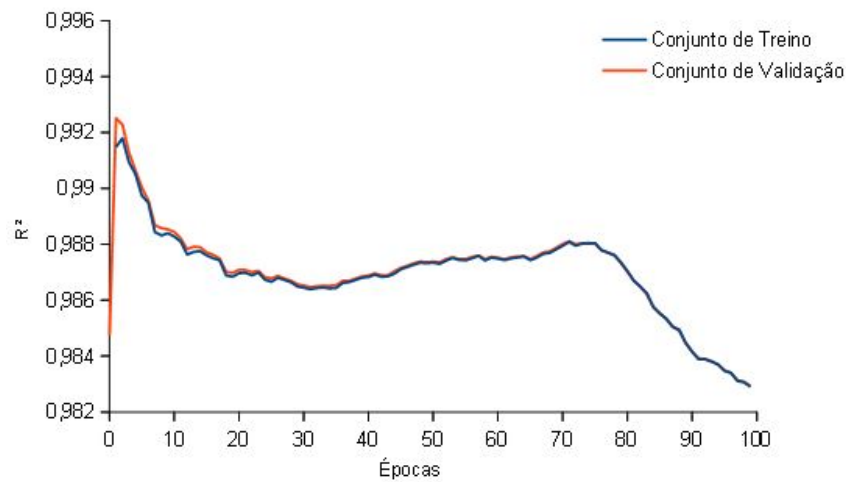
(c) Coeficiente de Determinação (R^2)

Figura 14 – Mean Square Error (MSE), Coeficiente de Pearson (R) e Coeficiente de Determinação (R^2) do período de treinamento e validação da rede recorrente.

Fonte: Autoria própria

Ao examinar o MSE, percebe-se que seu valor caiu constantemente até alcançar o número de 0,00004 no conjunto de treinamento enquanto no conjunto de validação houve um aumento considerável até finalizar com 0,024. Após alcançarem seus picos na terceira época ocorreu uma desvalorização tanto na correlação de Pearson que terminou o treinamento com 0,991 como no R^2 com 0,982 sendo os valores obtidos no treino iguais ao da validação para ambas as métricas.

Diferente da otimização do modelo neural de alimentação para frente, o gráfico do treinamento do modelo recorrente mostrou diversos sinais que pode ter ocorrido um *overfitting* ao decorrer das épocas, por exemplo, houve o aumento do MSE no conjunto de validação enquanto o seu par no conjunto de treino diminuía, outro sinal de *overfitting* está na variação da correlação de Pearson e R^2 que depois da terceira época começou a obter resultados piores até o final do treinamento.

Apesar da evolução desfavorável de suas métricas, um ponto que pode ser levado a favor da rede recorrente é que o mínimo local que ela alcançou foi muito baixo pois em todas as métricas ela iniciou com valores que não foram alcançados pela rede de alimentação para frente durante o decorrer de toda sua otimização.

4.5 ANÁLISE DAS PREDIÇÕES

Para realizar a predição no conjunto de teste referente a Novembro de 2019 foi selecionado o modelo com o peso de suas conexões que foi obtido na época em que a RNA de alimentação para frente e recorrente obtiveram a sua melhor pontuação no coeficiente de Pearson e R^2 que no caso foi a época 100 e 3 respectivamente.

Tabela 7 – Coeficientes de Pearson e R^2 obtido pelos preditores em suas predições

Preditor	Pearson	R^2
RNA - Recorrente	0,997	0,994
RNA - Alimentação Adiante	0,980	0,960
<i>Baseline</i>	0,980	0,960
Médias Móveis	0,282	0,080

Fonte: Autoria própria

A Tabela 7 permite comparar cada preditor ao analisar a correlação da sua predição com a realidade. O preditor composto pela rede recorrente foi o que fez as melhores predições, a segunda melhor predição foi obtida pelo preditor baseado em rede de alimentação para frente e pelo preditor *baseline* que empataram suas métricas. Dos quatro modelos o preditor baseado em médias móveis foi o que obteve o resultado negativo mais extremo pontuando 0,282 no coeficiente de Pearson e 0,080 no R^2 .

A hipótese que a rede recorrente sofreu de *overfitting* escrito na Seção 4.4 foi completamente equivocada, sendo que pelo ótimo resultado de suas predições no conjunto de teste indica que ela precisou de apenas 3 épocas em seu treinamento para abstrair em seus pesos a solução do problema. O modelo preditivo baseado em uma rede neural de alimentação para frente teve eficácia preditiva igual a dizer que o próximo minuto será igual ao minuto passado.

Médias móveis são utilizadas pela análise técnica dos preços como poderosas ferramentas de identificação da tendência atual do ativo pela sua capacidade de suavizar os movimentos dos preços, uma outra característica das médias móveis é que ela dá o mesmo peso a todos as cotações recebidas, independente se esta cotação esteja a 200 minutos ou a 10 minutos no passado no que resulta em uma predição atrasada, estas características são herdadas pelo preditor baseado em médias móveis o que causa uma correlação de Pearson e R^2 de valores muito a baixo do que os preditores baseados em redes neurais.

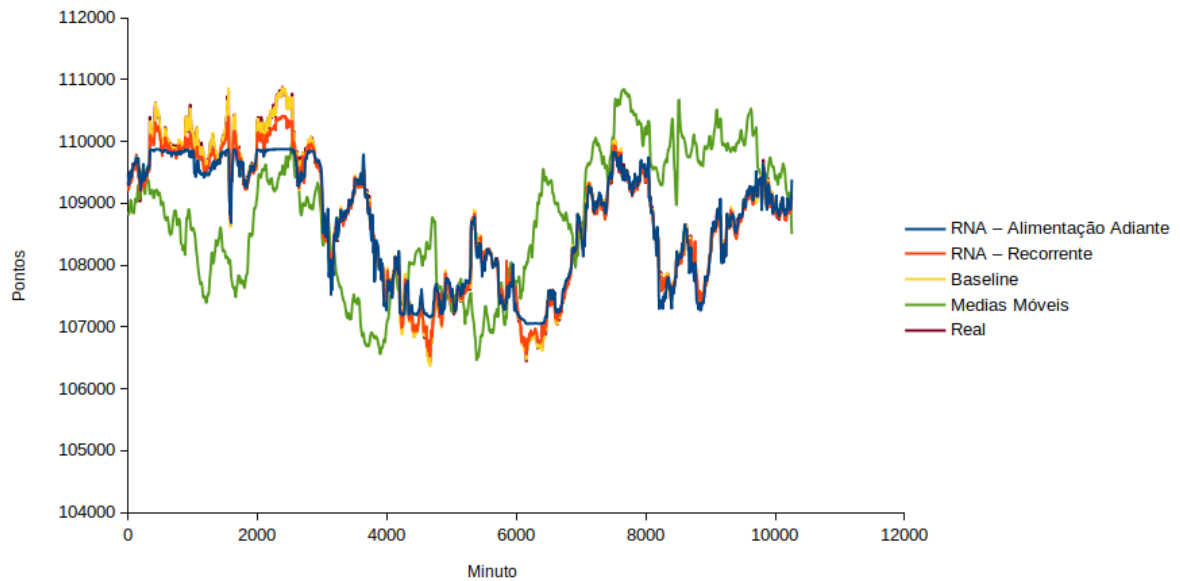


Figura 15 – Comparação entre as previsões das RNA avaliadas e os valores reais do conjunto de teste referente a novembro de 2019.

Fonte: Autoria própria

A Figura 15 mostra o gráfico com as previsões obtidas pelos modelos listados na Tabela 7 em contraste com o real valor cotado. Percebe-se que no gráfico que a linha destinada à cotação real é constantemente sobrescrita pelas previsões dos preditores baseados em redes neurais e pelo baseline. é notável também o grande afastamento que há entre o preço real e o preditor baseado em médias móveis.

Outro importante fato a ser observado foi a capacidade dos modelos baseados em redes neurais conseguirem fazer boas previsões quando o mercado de contrato do mini Índice Bovespa estava sem uma tendência principal definida diferente do que foi mostrado a ela durante o treinamento (Imagem 12a) e validação (Imagem 12b) que exhibe dois gráficos com tendência incontestavelmente de alta.

4.6 SIMULAÇÃO DE NEGOCIAÇÃO NO MERCADO DE CONTRATO FUTURO DO MINI ÍNDICE BOVESPA EM DEZEMBRO DE 2019

É de interesse para este trabalho saber se um robô de negociação que faz uso do melhor modelo preditivo obtido da Seção 4.5 obtém uma vantagem significativa na negociação dos contratos de futuro do mini Índice Bovespa. Para cumprir este propósito, foi realizada uma experiência, um teste empírico simulando um agente de negociações no mercado em dezembro de 2019 por um agente que adota as seguintes regras para negociar:

- O agente pode manter apenas uma posição no mercado.
- O agente se posiciona como vendido toda vez que um preço menor do que o atual for predito para o próximo minuto.
- O agente se posiciona como comprado toda vez que um preço maior do que o atual for predito para o próximo minuto.

Para fins de praticidade, não foram computadas as taxas e impostos que são comuns para esta atividade, além disso foram usados pontos percentuais como unidade de medida para o saldo de negociação. O resultado da simulação pode ser visto na Figura 16 que exhibe o gráfico da evolução do saldo ao decorrer do mês enquanto o agente realizava as suas negociações no mercado.

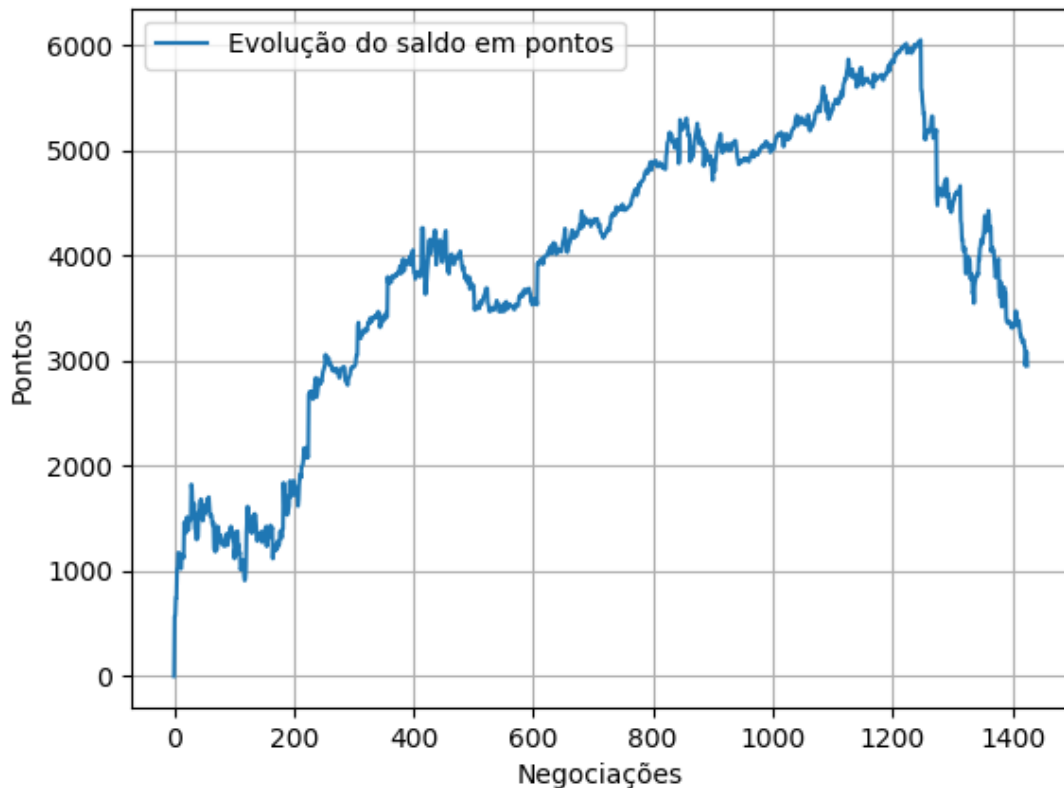


Figura 16 – Evolução do saldo de pontos obtidos pelo robô em dezembro de 2019.

Fonte: Autoria própria

O gráfico da Figura 16 deixa evidente que houve uma maior quantidade de negociações lucrativas por causa da tendência positiva na evolução do saldo, o agente acumulou seu maior valor até fechar 1.200 transações conseguindo juntar 6.055 pontos. Após este pico, ocorreu uma sucessão de negociações com resultados negativos até terminar com 2.955 pontos após 1.426 transações.

Apesar de ter obtido lucro, o desempenho do robô foi pior que a valorização da bolsa de valores no período, pois o mercado abriu dezembro precificando o contrato de Índice Bovespa mini em 108.315 pontos e no final do mês após uma forte tendência de alta fechou o preço do contrato em 115.770 pontos o que resultou em um superavit de 7.455 superando o resultado do robô em 4.500 pontos.

O gráfico da Figura 17 deixa a entender que o desempenho inferior do agente em relação ao mercado se deve a forte apreciação no valor do Índice Bovespa neste período e como o robô muda sua posição de comprado para vendido toda vez que prediz uma tendência de desvalorização acabou perdendo pontos por desconfiar do fim da tendência de alta em diversas

ocasiões ao decorrer do mês.

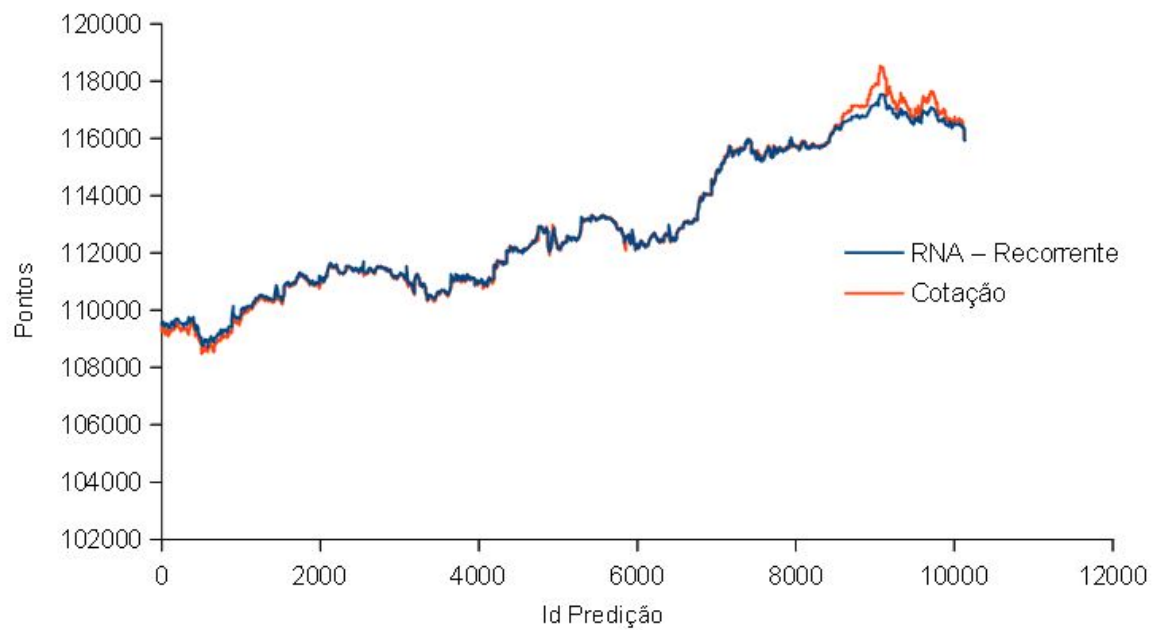


Figura 17 – Predição da rede de alimentação recorrente pareada com a cotação de dezembro de 2019 do contrato futuro do mini Índice Bovespa.

Fonte: Autoria própria

5 CONCLUSÕES

Na Seção 1.3, foram definidos os objetivos específicos para serem concluídos e assim alcançar o objetivo geral deste trabalho o qual é realizar um comparativo das predições realizadas por uma RNA de conexões recorrente com uma outra RNA que usa de alimentação para frente no preço do contrato futuro do mini Índice Bovespa.

No final dos experimentos a rede de alimentação recorrente foi o modelo com as melhores predições com uma correlação de Pearson em 0.997. Ao fazer um comparativo sobre a predição da RNA de alimentação recorrente com a de outros modelos regressivos listado na Tabela 7 da Seção 4.5 foi constatado a sua superioridade com as melhores métricas.

Apesar de não ser o melhor preditor neste trabalho a RNA de alimentação adiante alcançou uma correlação de Pearson em 0.98, ficando apenas 1,7 ponto percentual abaixo do modelo de alimentação recorrente. Em comparação com os outros modelos regressivos a RNA de alimentação adiante também se saiu bem, fez predições superiores ou iguais à linha de base, superando o preditor baseado em médias móveis e empatando com as predições do modelo que prevê a imutabilidade do preço.

Com o interesse em saber como seria o desempenho de um robô que faz uso do melhor modelo regressivo como conselheiro para suas tomadas de ações a Seção 4.6 fechou os experimentos desta monografia com uma simulação de negociações no mercado de contratos futuros do mini Índice Bovespa conseguindo obter lucro no final de dezembro em 2019, porém não superou o retorno obtido naturalmente se estivesse optado por uma carteira de ações baseado no Índice Bovespa, pois o houve uma forte tendência de alta neste mês e o agente ao alternar entre comprado e vendido acabou por não aproveitar ao máximo o período de alta.

5.1 TRABALHOS FUTUROS

Após a conclusão desta monografia, segue em aberto as seguintes linhas de pesquisa que podem dar continuidade a este trabalho:

- Usar taxa de juros básica como um parâmetro de entrada.
- Replicar o comparativo entre RNAs de alimentação recorrente para taxas de câmbio pois este é um mercado que tem os fundamentos diferentes dos mercados de ações.
- Realizar o comparativo entre RNAs de alimentação recorrente e para frente em uma sazonalidade diferente da deste trabalho.
- Realizar o comparativo entre RNAs de alimentação recorrente e para frente em ações com baixo volume de negociações.
- Fazer um estudo da predição do mercado de contratos futuros para o mini Índice Bovespa usando uma rede que possui camadas de alimentação para frente e recorrentes em sua arquitetura.
- Testar diferentes estratégias na bolsa de valores com base nas predições da RNA com alimentação recorrente.

REFERÊNCIAS

- ADEBIYI, A. A. et al. Stock Price Prediction using Neural Network with Hybridized Market Indicators. 2016. Disponível em: <http://eprints.covenantuniversity.edu.ng/4112/1/Emerging_Trend.pdf>.
- ASSAF, N. **Mercado Financeiro**. São Paulo: Atlas, 2014.
- B3. **Futuro mini de Ibovespa**. 2018. Disponível em: <http://www.b3.com.br/pt_br/produtos-e-servicos/negociacao/renda-variavel/futuro-mini-de-ibovespa.htm>. Acesso em: 16 de outubro de 2018.
- BEIL, J.; PERNER, G.; ASFOUR, T. Design and control of the lower limb exoskeleton KIT-EXO-1. **IEEE International Conference on Rehabilitation Robotics**, v. 2015-Septe, p. 119–124, 2015. ISSN 19457901.
- BRAGA, A. d. P.; CARVALHO, A. P. d. L. F. d.; LURDEMIR, T. B. **Redes Neurais Artificiais Teoria e Aplicações**. Rio de Janeiro: LTC Livros Técnicos e Científicos, 2000.
- CHUNG et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. 2014. Disponível em: <<https://arxiv.org/abs/1412.3555>>.
- COLAH, C. **Understanding LSTM Networks**. 2011. Disponível em: <<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>>. Acesso em: 25 de novembro de 2018.
- COPPIN, B. **Inteligência Artificial**. Rio de Janeiro: LTC, Livros Técnicos e Científicos, 2010.
- FACELI, K. et al. **Inteligência Artificial: Uma abordagem de Aprendizado de Máquina**. Rio de Janeiro: LTC Livros Técnicos e Científicos, 2011.
- FAYYAD, U.; GRINSTEIN, G.; WIERSE, A. **Information Visualization in Data Mining and Knowledge Discovery**. [S.l.]: Morgan Kaufmann, 2001.
- FISCHER, A. et al. **Mercado de Valores Mobiliários Brasileiro**. [s.n.], 2014. 376 p. ISBN 978-85-67896-00-7. Disponível em: <www.cvm.gov.br/menu/investidor/publicacoes/livros.html>.
- FORTUNA, E. **Mercado Financeiro Produtos e Serviços**. São Paulo: QUALITYMARK, 1992.
- GIACOMEL, F. d. S. Um Método Algorítmico para Operações na Bolsa de Valores Baseado em Ensembles de Redes Neurais para Modelar e Prever os Movimentos dos Mercados de Ações. 2016. Disponível em: <<https://www.lume.ufrgs.br/bitstream/handle/10183/134586/000988403.pdf?sequence=1>>.
- GOYAL, P.; PANDEY, S.; JAIN, K. **Deep Learning for Natural Language Processing Creating Neural Networks with Python**. Bangalore, Karnataka, India: LTC Livros Técnicos e Científicos, 2018.

- HAYKIN, S. **Redes Neurais Princípios e prática**. São Paulo: ARTMED EDITORA S.A, 2001.
- HEBB, D. **The Organization of Behavior: A Neuropsychological Theory**. [S.l.]: Psychology Press, 1949.
- HOCHREITER, S.; SCHMIDHUBER, J. *Ltsm*. v. 9, n. 8, p. 1–32, 1997. ISSN 0899-7667.
- HOPFIELD, J. J. **Neural networks and physical systems with emergent collective computational abilities**. New Jersey: Proc. Natl. Acad. Sci. USA 79, 1982.
- JUNQUEIRA, L.; CARNEIRO, J. **Histologia Básica**. Rio de Janeiro: EDITORA GUANÁBARA KOOGAN S.A, 2008.
- KUNIAKI, U. et al. Deep learning for stock prediction using numerical and textual information. 2016. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/7550882>>.
- LINHARES, S.; GEWANDSZNAJDER, F. **Biologia Hoje - Volume 1 (Citologia / Histologia / Origem da Vida)**. São Paulo: Atica, 2006.
- MANASWI, N. K. RNN and LSTM. In: Deep Learning with Applications Using Python. 2018. Disponível em: <https://doi.org/10.1007/978-1-4842-3516-4_9>.
- MEDSKER, L.; JAIN, L. **Recurrent Neural Networks Design and Applications**. [S.l.]: CRC Press, 2001.
- MENDEL, L.; GALVÃO, A. M.; OLIVEIRA, V. I. **Mercado financeiro: Uma Abordagem Prática dos Principais Produtos e Serviços**. São Paulo: Fleuriet, Michel, Galvão, Alexandre, Mendes Lucass, 2006.
- MENDES, M. **TECIDO NERVOSO (E MAIS!)**. 2013. Disponível em: <http://maxaug.blogspot.com/2013_11_01_archive.html>. Acesso em: 17 de outubro de 2018.
- MENESES, A.; MARIANO, F. **Mercado Financeiro: Teoria, Questões Comentadas E Mais De 100 Questões Propostas**. Rio de Janeiro: Elsevier Editora Ltda, 2011.
- MINSKY, M.; PAPERT, S. **Perceptrons: An introduction to computational geometry**. Cambridge: MIT Press, 1969.
- NELSON, D. M. Q. USO DE REDES NEURAIAS RECORRENTES PARA PREVISÃO DE SÉRIES TEMPORAIS FINANCEIRAS. 2017. Disponível em: <<https://www.dcc.ufmg.br/pos/cursos/defesas/1999M.PDF>>.
- ROSA, J. L. G. **Fundamentos da inteligência artificial/**. Rio De Janeiro: LTC, 2011.
- SCHOSSLER, J. L. PREVISÃO DE VALOR DE UM ATIVO FINANCEIRO UTILIZANDO REDES NEURAIAS ARTIFICIAIS PARA PROCESSAMENTO DE DADOS DE SÉRIES TEMPORAIS. 2017. Disponível em: <<http://repositorio.roca.utfpr.edu.br/jspui/handle/1/14431>>.
- SILVA, C. d.; SASSOM, S. **Biologia 1**. São Paulo: Saraiva, 2005.

SUNDERMEYER, M.; NEY, H.; SCHLÜTER, R. From Feedforward to Recurrent LSTM Neural Networks for Language Modeling. 2015. Disponível em: <<https://ieeexplore.ieee.org/document/7050391>>.

TOSCANO, L. C. **Guia de referencia para o MERCADO FINANCEIRO**. São Paulo: Editco Comercial Ltda, 2004.

WIDROW, B.; HOFF, T. **An adaptive "ADALINE"neuron using chemical "memistors"**. [S.l.: s.n.], 1960.

ZAMBIASI, S. P. **O Neurônio Artificial**. 2018. Disponível em: <https://www.gsigma.ufsc.br/~popov/aulas/rna/neuronio_artificial/index.html>. Acesso em: 17 de outubro de 2018.