

# Métodos de *Deep Learning* Aplicados a *Candlestick* como Estratégia de Investimento

Felipe Dias Paiva

Centro Federal de Educação Tecnológica de Minas Gerais – CEFET/ MG  
fpaiva@dcsa.cefetmg.br

Carolina Magda da Silva Roma

Universidade Federal de Lavras – UFLA  
carolina.roma@ufla.br

## RESUMO

O objetivo desta pesquisa foi investigar se métodos de *deep learning*, representados pelo *Convolutional AutoEncoder* (CAE) e *Long Short-Term Memory* (LSTM), são capazes de em conjunto prever o movimento do preço de ações negociadas no mercado acionário brasileiro, caracterizado como um mercado emergente, baseando-se no desempenho *out-of-sample* como métrica de avaliação. Para tanto, utilizamos como objeto de investigação as ações preferenciais da empresa Petrobrás. Os resultados apresentados apontam que a técnica de CAE alcançou uma acurácia de aproximadamente de 90% no processo de encoder dos gráficos de *candlesticks*. Essa compactação da imagem serviu de insumo para aplicação da rede neural LSTM que conseguiu resultados regressores superior a um modelo baseado na técnica de *support vector machine*. Além de apresentar um bom potencial de classificação, com *precision* de 73%, métrica superior à alcançada por decisões de caráter naive.

**Palavras-chave:** *Deep Learning*, *Convolutional AutoEncoder*, *Long Short-Term Memory*, Mercado Financeiro, *Candlesticks*.

## 1 INTRODUÇÃO

A Hipótese de Eficiência de Mercado, apresentada por Fama (1970), ganhador do prêmio nobel de Economia em 2013, e fundamento da Moderna Teoria de Finanças, sintetiza que um mercado eficiente o preço do ativo reflete toda e qualquer informação disponível. Assim, a análise histórica do comportamento dos preços não geraria oportunidades de ganho a longo prazo para os investidores, uma vez que, existindo ineficiências de mercado, as mesmas seriam exploradas pelos participantes e eliminadas. Contudo, os mercados tendem a não ser eficientes e são considerados, de forma geral, como caóticos, complexos, voláteis e dinâmicos (Singh; Srivastava, 2016). Neste sentido, a análise técnica que estuda os movimentos históricos dos preços visando identificar oportunidades de investimento vantajosas ganha destaque juntamente com um corpo de literatura que visa através da aplicação de métodos estatísticos tradicionais ou, recentemente, de técnicas mais sofisticadas de inteligência artificial prever o comportamento de mercado.

Na área de Finanças, a abordagem de máquinas de aprendizado, que leva em consideração a relação não linear entre as variáveis, tais como redes neurais artificiais (*artificial neural network*, ANN), máquinas de suporte de vetor (*support vector machines*, SVM), *random forest* (RF), têm sido testadas na previsão de movimento da carteira de mercado (Kim, 2003), falência bancária (Le; Viviani, 2018), *credit ratings* (Ozturk; Namli; Erdal, 2016), dentre outros tópicos. As aplicações nessa metodologia podem ser

empregadas usando experimentos tanto supervisionados quanto não supervisionados, que se diferenciam pela presença ou não dos rótulos nos dados.

Nos anos mais recentes, o método de *deep learning*, representado, por exemplo, pela *Convolutional Neural Network* (CNN) – rede neural supervisionada cuja proposta é de classificar os dados - e *Convolutional AutoEncoder* (CAE) – rede neural auto-supervisionada que busca reduzir a dimensionalidade - e *Recurrent Neural Networks* (RNN), tem sido amplamente adotado nas atividades de processamento de linguagem natural (Kiros; Salakhutdinov; Zemel, 2014) e reconhecimento de voz (Graves; Mohamed; Hinton, 2013). Especificamente, Lecun, Bengio e Hinton (2015, p. 436) apontam que *deep learning* permite que a representação dos dados com múltiplos níveis de abstração seja aprendida por modelos que tenham múltiplas camadas de processamento e, ainda mais, avançaram significativamente o estado da arte em reconhecimento da fala, reconhecimento de objeto visual, detecção de objetos, dentre outros. Para usar essa abordagem, usa-se a conversão dos dados brutos em uma imagem 2D ou, então, a utilização a partir do vetor de sinal 1D onde, por exemplo em finanças, Chen e He (2018) empregam essa última no estudo da previsão do movimento do preço da ação no mercado chinês, enquanto que Velay e Daniel (2018) usam as informações do preço de abertura, máximo, mínimo e fechamento do ativo na forma de *candlestick* para treinar a CNN na detecção de padrões comuns nos gráficos. É relevante destacar que estudos de *deep learning* com séries financeiras ainda são escassos (Hosaka, 2019) e, principalmente, quando aplicados em ativos de mercados emergentes que podem apresentar um desempenho diferente das evidências destacadas em países desenvolvidos.

Neste sentido, a pesquisa aqui delineada propõe-se a aprofundar a literatura especializada ao propor um modelo híbrido a partir de dois métodos de *deep learning*, que são a CAE e a *Long Short-Term Memory* (LSTM) na construção de uma estratégia de investimento a partir da transformação dos preços de abertura, máximo, mínimo e fechamento diário em *candlesticks* que demonstram a evolução do comportamento histórico do ativo. A arquitetura da CAE é utilizada no processamento da imagem e busca armazenar o máximo possível as características principais do dado de input, de forma que se possa reconstruir o objeto em caminho reverso. Já a LSTM é um tipo de RNN que extrai a informação dependente de longo prazo no processo de previsão e permite lidar com o problema de desaparecimento do gradiente (Sak; Senior; Beaufays, 2014), o qual pode acontecer na RNN. A mesma é utilizada em principalmente em situações que envolvam séries temporais.

As principais contribuições desta pesquisa estão i) na investigação de métodos de *deep learning*, em específico o CAE e LSTM em conjunto, visto que poucos são os trabalhos nessa linha. Em específico, o estudo de Hu et al.

(2018) se aproxima a proposta corrente no sentido de usar CAE para explorar estratégias de investimento, contudo, aqui distancia-se com a proposição de um metodologia híbrida; ii) no fornecimento de novas evidências acerca do desempenho de métodos de *deep learning* para previsão dos preços com dados de um país emergente e não asiático, cujas pesquisas nessa área ainda são incipientes.

O artigo é estruturado em cinco seções. A primeira compreende a introdução, seguida por uma breve revisão de pesquisas relacionadas acerca de CNN, CAE e LSTM. Na terceira seção, os aspectos metodológicos são apresentados. A quarta seção apresenta e discute os resultados encontrados e, por último, as considerações gerais são realizadas.

## 2 REVISÃO DA LITERATURA

### 2. 1 Convolutional Neural Network e Convolutional AutoEncoder

*Deep learning* é um método capaz de gerar uma maior característica (*feature*) dimensional, resultando em um melhor desempenho do modelo. Desta forma, A CNN, caracterizada como um método de *deep learning* e rede *feedforward* que explora a extração de características e mais adequada para aplicação em imagens, não é relativamente nova, contudo, somente nos anos mais recentes é que recebeu maior aplicação sendo uma abordagem reconhecida em atividades de visão computacional, conforme destacam Krizhevsky, Sutskever e Hinton (2012). Os autores implementaram a CNN para classificação de 1,2 milhão de imagens com alta resolução no banco de dados da ImageNet alcançando resultados mais satisfatórios que os previamente reportados, com uma taxa de erro top-1 e top-5 de 37,5% e 17%, respectivamente.

O uso dessa metodologia com séries financeiras também demorou um pouco para ganhar espaço com sua utilização em classificação de imagens. Stojanovski et al. (2015) emprega a arquitetura de *deep CNN* na análise de sentimento a partir de mensagens no Twitter. Ding et al. (2015) integram a metodologia de *event embedding* com uma *deep CNN* no estudo da influência de eventos financeiros extraídos da Reuters e Bloomberg no movimento do preço da ação e S&P500. Hosaka (2019) propõe a transformação de índices financeiros extraídos do balanço patrimonial e declaração de lucros e perdas de empresas de bolsas de valores no Japão em formato de imagem como input para CNN e encontram que a mesma é capaz de prever falência, como também, o desempenho desse modelo é superior a outras abordagens, tais como análise discriminante linear, SVM e multi-layer perceptrons (MLP).

CAE também é um método de *deep learning* utilizado para reduzir a dimensionalidade que também tem sido implementado com dados financeiros. Neste sentido, Hu et al. (2018), por exemplo, apontam que uma estratégia de investimento que alimenta a CAE com imagens de *candlestick* para determinar características profundas (*deep features*) como representação do ativo, em seguida, obtêm as similaridades das ações a partir de tais características para clusterizar os ativos, o que é realizado através da otimização modular e, por último, seleciona um ativo por cluster com base no maior índice de Sharpe (1994) para composição de um portfólio resulta em um retorno total superior a uma

estratégia que substitui a parte do CAE pelos dados brutos (283,5% e 208,8%, respectivamente), como também, supera o retorno do índice *Financial Times Stock Exchange 100 Index* (FTSE 100) e outros fundos.

### 2.2 Long Short-Term Memory (LSTM)

A RNN é um tipo de ANN, delineada para assimilar padrões sequenciais ou variantes no tempo (Medsker; Jain, 2001; Gao; Shai, 2018), enquanto que LSTM é uma variante da RNN e encontra-se entre os mais avançados métodos de *deep learning* aplicados a aprendizagem sequencial (Fischer; Krauss, 2018). Na área financeira ambas têm sido utilizadas para incorporar estados passados e com diferentes aplicações. Gao e Shai (2018) utilizam a LSTM junto com informações de negociação do ativo (abertura, máximo, mínimo, fechamento, preço ajustado e volume) e 15 indicadores de análise técnica aplicando a técnica análise de componentes principais sobre esses últimos, buscando considerar a correlação entre os mesmos, com o objetivo de prever o preço de fechamento do dia seguinte. Os resultados destacam a superioridade do modelo proposto em relação ao SVM, ARMA e GARCH, dentre outros. Lv et al. (2018) focam na LSTM juntamente com a otimização de enxame de partículas para previsão de preço.

Fischer e Krauss (2018) empregam a LSTM em um amplo conjunto de ações ao investigar a sua capacidade de previsão usando todos os ativos constituintes do índice S&P 500 de 1992 a 2015, contrastando a metodologia com benchmarkings representados pelo RF, uma rede neural profunda e regressão logística. Velay e Daniel (2018) investigam se CNN e LSTM são capazes de identificar padrões comuns em gráficos de *candlestick* (representando o preço de abertura, máximo, mínimo e fechamento) e de linha na ação Alphabet C. A arquitetura da CNN segue a AlexNet, com camadas de convolução conectadas seguidas por camadas completamente conectadas, e na LSTM os nós centrais representam um estado que recebe atualizações via diferentes portões sendo utilizada para gerar o output, conforme descrevem os autores. Os dois métodos de *deep learning* foram comparados com um detector codificado construído e a LSTM obteve uma taxa de recall de 97%, sendo superior a 2D CNN (73%) e 1D CNN (64%). Contudo, aplicando a abordagem para dados de outra companhia não foi possível alcançar um bom desempenho alertando a necessidade do comportamento histórico de múltiplos ativos serem levados em conjunto no treinamento do modelo.

## 3 DESIGN DO EXPERIMENTO

Para a elaboração da presente pesquisa, definiu-se por utilizar as ações preferenciais da empresa Petróleo Brasileiro S/A (Petrobrás – PETR4). A escolha pela Petrobrás como objeto de estudo, deveu-se ao fato de suas ações possuírem alta liquidez e representatividade frente ao mercado de capitais brasileiro. Portanto, trabalhou-se com uma amostra não probabilística intencional, mantendo como principal fonte de dados as séries temporais dos preços de abertura, máximo, mínimo e fechamento do referido ativo no período de 02/01/2001 a 29/09/2016. As referidas cotações diárias da PETR4 foram extraídas já ajustadas para

divididos, bonificações, *split* e *implit* das ações. A coleta de dados deu-se por meio do terminal Bloomberg.

Com relação ao modelo proposto, apresentamos por meio da Figura 01 um esquema que representa o

experimento delineado e testado. Destacamos quatro grandes etapas delineadas no experimento, que podem ser visualizadas na Figura 01 e que são devidamente detalhadas no decorrer da seção.

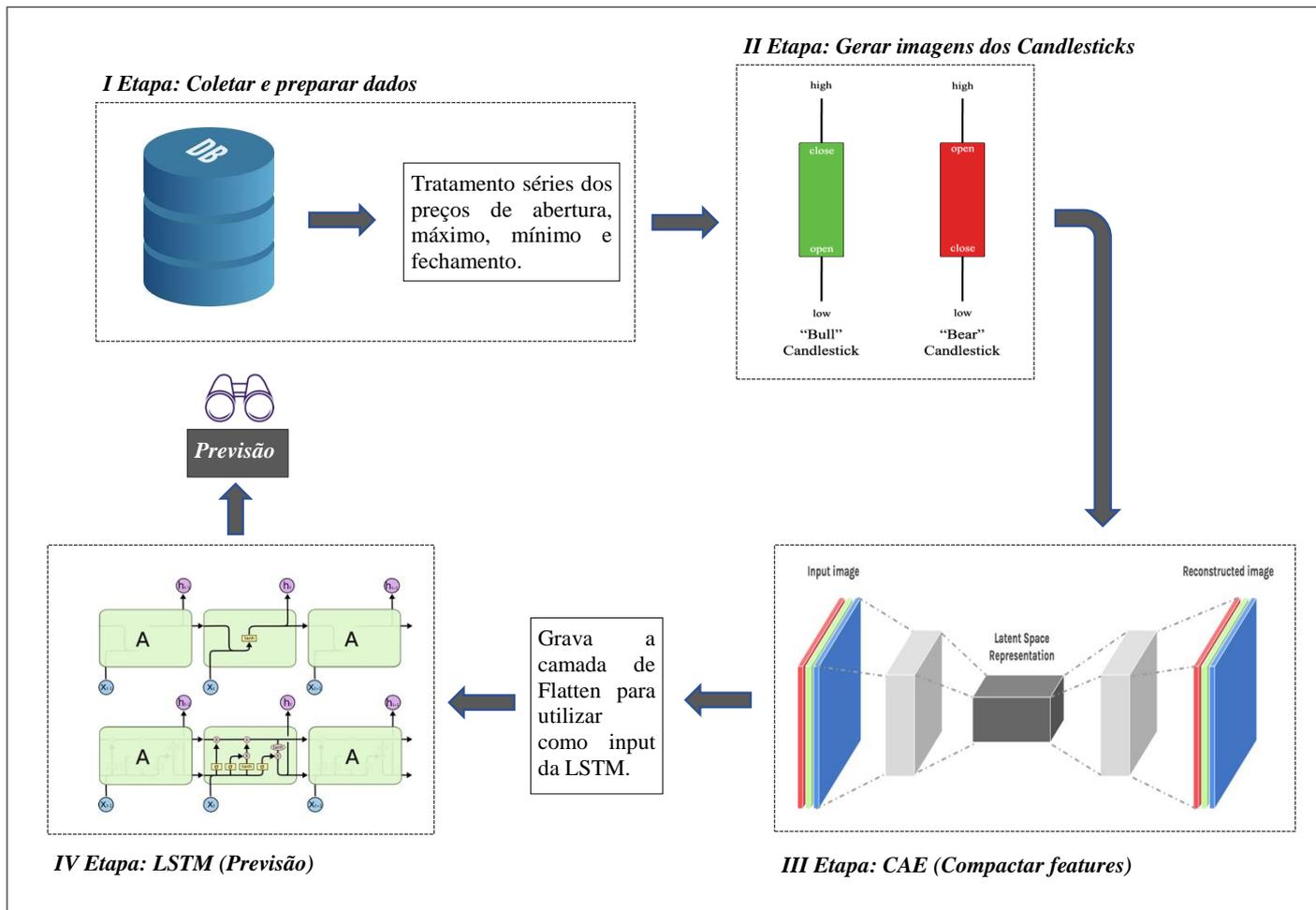


Figura 01- Esquema do experimento

### 3.1 Etapas do modelo proposto

#### 3.1.1 Coleta e preparo dos dados

Dado que definiu-se por utilizar gráficos de Candlesticks Heikin-Ashi (HA). A primeira etapa de destaque do algoritmo foi calcular as variáveis que compõem o gráfico. Os quatro pontos necessários para formação do HA são assim estabelecidos:

$$HA\_Fech_t = (Abert_t + Max_t + Min_t + Fech_t) / 4$$

$$HA\_Abert_t = (HA\_Abert_{t-1} + HA\_Fech_t) / 2$$

$$HA\_Max_t = \text{Maior}(Max_t, HA\_Abert_t, HA\_Fech_t)$$

$$HA\_Min_t = \text{Menor}(Min_t, HA\_Abert_t, HA\_Fech_t)$$

Onde:

HA\_Fech<sub>t</sub> : preço de fechamento utilizado na formação do HA no período *t*

HA\_Abert<sub>t</sub> : preço de abertura utilizado na formação do HA no período *t*

HA\_Min<sub>t</sub> : preço mínimo utilizado na formação do HA no período *t*

HA\_Max<sub>t</sub> : preço máximo utilizado na formação do HA no período *t*

Abert<sub>t</sub> : preço de abertura do candle tradicional no período *t*

Max<sub>t</sub> : preço máximo do candle tradicional no período *t*

Min<sub>t</sub> : preço mínimo do candle tradicional no período *t*

Fech<sub>t</sub> : preço de fechamento do candle tradicional no período *t*

#### 3.1.2 Geração do Gráficos dos Candlesticks Heikin-Ashi

Preparados os dados para construção dos HA, a segunda etapa consistiu na geração das imagens e arquivamento das mesmas para utilização na etapa seguinte. Após testes para horizontes temporais diferentes, definimos por imagens de HA para 5 dias. A Figura 02 ilustra um exemplo das 3.900 imagens geradas. As imagens foram geradas em formato de cor RGB, tamanho 128 x 128 pixels, ou seja, para cada imagem temos 16.384 pontos para cada uma das escalas de cor.

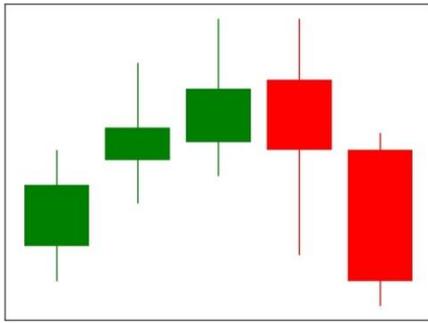


Figura 02- Exemplo de imagem gerada do *candlestick Heikin-Ashi*

Ainda sobre a Figura 02, esclarecemos que os candlesticks na cor verde indicam que o preço de abertura é menor que o de fechamento, já os candles vermelho têm relação oposta, ou seja, preço de abertura maior que o preço de fechamento.

### 3.1.3 Convolutional AutoEncoder dos Candlesticks

Geradas e arquivadas as imagens dos HA, inicia-se a terceira etapa que foi gerar a *Convolutional AutoEncoder* (CAE) com a finalidade de redução de dimensionalidade das *features* a serem utilizadas na modelagem de previsão. Definimos para o presente trabalho estruturar a etapa de treinamento da CAE com 3.800 imagens e de teste com 100 imagens. Destacamos na Figura 03 a configuração da CAE que utilizamos no presente trabalho.

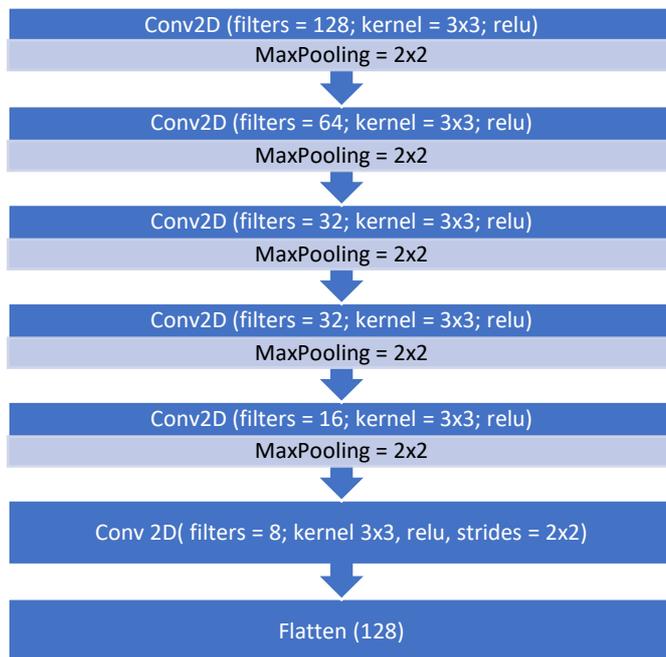


Figura 03 – Arquitetura CAE

### 3.1.4 Long Short-Term Memory para Previsão dos Retornos

Imagens vetorizadas deu-se início a etapa final de previsão do retorno máximo que um investidor poderia alcançar nos próximos 3 dias. Sendo que, esse retorno máximo foi calculado estabelecendo que o investidor compra o ativo no preço de abertura do dia seguinte a formação da imagem e estaria hábil a comercializar o papel

no preço máximo de qualquer um dos próximos 3 dias da formação do HA. Para fins de previsão nas fases de treinamento e teste, verificou-se o maior valor entre o preço máximo desses próximos 3 dias da formação do HA, este então foi utilizado para cálculo do retorno máximo a ser previsto. A técnica utilizada para previsão do retorno foi a LSTM utilizando como input as *features* geradas pelo CAE. A configuração da LSTM utilizada na pesquisa pode ser visualizada na Figura 04.

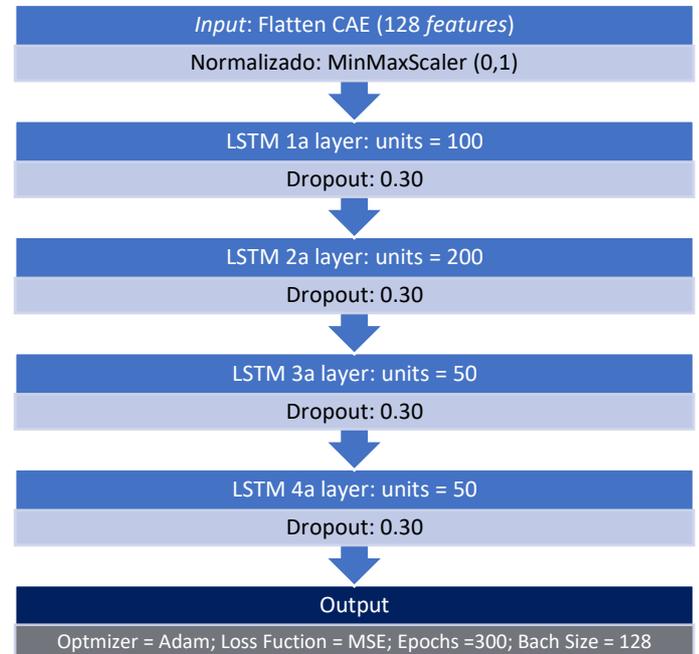


Figura 04 – Arquitetura LSTM

### 3.2 Métricas de Desempenho

Para análise de desempenho de diferentes etapas e propósitos utilizamos as seguintes medidas: a partir da matriz de confusão (ou tabela de contingência), a qual representa as possíveis combinações dos resultados estimados em relação aos valores reais, computou-se as métricas de *acurácia* e *precision*. Também foi calculada a raiz do erro quadrático médio como métrica de *acurácia* do modelo. As formulações matemáticas para cálculo das métricas estão dispostas na Tabela 01.

Tabela 01 – Métricas de Desempenho

Acurácia <sup>a</sup>	Precision <sup>a</sup>	RMSE <sup>b</sup>
$\frac{VP + VN}{VP + VN + FP + FN}$	$\frac{VP}{VP + FP}$	$\sqrt{\frac{\sum_{i=1}^n (A_i - P_i)^2}{n}}$

<sup>a</sup> Em uma classificação binária como utilizado neste trabalho, têm-se as seguintes opções: verdadeiro positivo (VP), verdadeiro negativo (VN), falso positivo (FP) e falso negativo (FN). <sup>b</sup> Sendo que  $A_i, P_i$  e  $n$  representam, respectivamente, o retorno atual ou realizado, o retorno estimado e o tamanho do conjunto de dados.

### 3.3 Baseline

Para comparação do desempenho do modelo proposto baseado em *deep learning*, utilizaremos como

*baseline* uma técnica de *machine learning* que tem alcançados bons resultados quando aplicada a problemas da área financeira que é a técnica *support vector machine* (SVM), além da tomada de decisão *naive*.

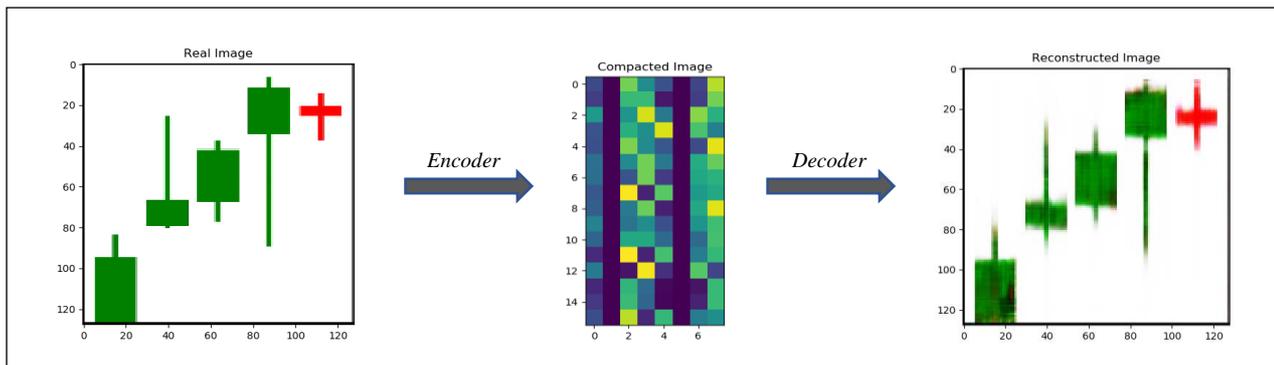
### 3.4 Recurso computacional

O algoritmo utilizado na pesquisa foi escrito na linguagem de programação *Python*. Como recurso computacional para processamento do algoritmo utilizamos uma Workstation: dual Xeon com 40 núcleos, 64 GB de memória RAM e GPU GeForce GTX 1080 ti.

## 4 RESULTADOS

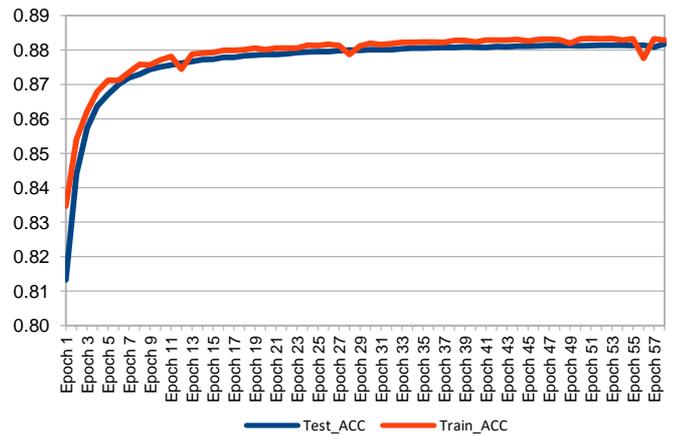
Tal como explicitado no capítulo anterior, iniciamos a execução do experimento calculando as variáveis para construção dos preparando as séries temporais dos *Candlesticks* Heikin-Ashi. Preparado as variáveis, executamos a construção dos imagens dos HA, com representação de 5 dias. Foram geradas 3.900 imagens em 3 canais de cores. Posteriormente as imagens geradas foram utilizados no modelo CAE, com fins a compactar as *features* da imagem dos HA. A Figura 05 expõe aleatoriamente uma das imagens que passou pela técnica.

Podemos por meio da Figura 05 observar a imagem real, que foi gerada para o tamanho de 128x128 pixels, ou seja, a vetorização da mesma implicaria num vetor com 16.384 *features*. Após o processo de encoder da imagem, a mesma foi reduzida para um vetor de 128 *features*, que na Figura 05 a título de ilustração é apresentada numa matriz 16x8. Como recorte final, é exposto a reconstrução da imagem inicial tendo como ponto de partida a imagem compactada. Numa comparação visual, podemos observar grande similaridade entre a imagem real e a imagem reconstruída a partir das variáveis compactadas.



**Figura 05- Exemplo do processo de encoder e decoder baseado na técnica de convolucional**

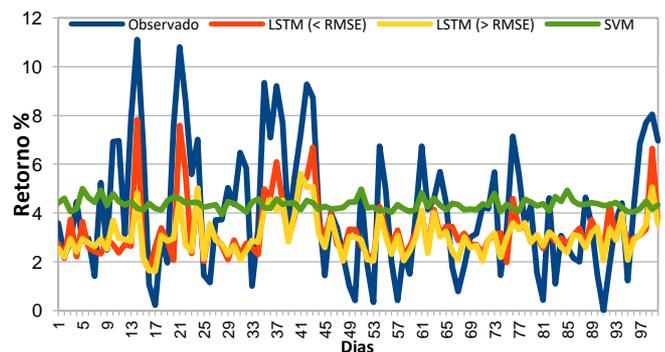
Para uma melhor explicação da eficácia de uso da técnica de compactação, apresentamos a Figura 06 que indica que a acurácia média da imagens reconstruídas face a imagens reais foi de 88%, tendo convergido em 57 épocas. Esses números tratam-se de uma média obtida entre as 50 iterações realizadas. Verificamos também uma grande proximidade de desempenho dos resultados dos conjuntos de treinamento e teste.



**Figura 06- Medida de acuracidade CAE**

Realizado o processo de encoder das imagens, utilizamos o mesmo como input para a rede neural LSTM. Após diversos testes, acabamos decidindo por utilizar um *lag* de 90 imagens. Assim, para a fase de treinamento passamos a ter 3.710 instâncias, para o conjunto de teste foram mantidas as 100 instâncias. Dado que se tratar também de uma técnica estocástica, realizamos novamente 50 iterações. Sendo que, para apresentação dos resultados alcançados plotamos na Figura 07 a previsão das iterações da LSTM que obtiveram maior e menor RMSE, além dos retornos observados e do *baseline* calculado a partir da técnica de *support vector machine*.

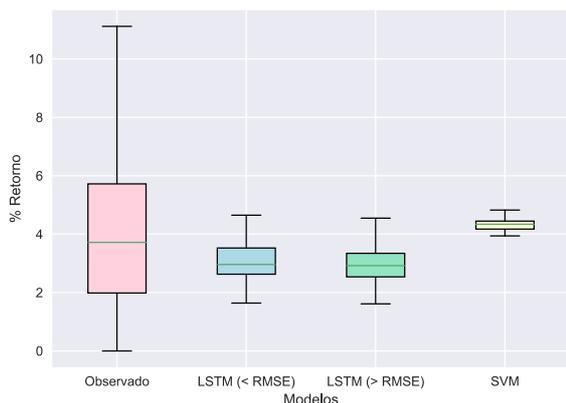
A título de análise gráfica, notamos um comportamento *flat* da previsão da SVM e um comportamento mais oscilatório do modelo LSTM, que acompanha melhor a série de retornos reais, mas que não segue com proximidade os picos de variação da série observada. As deduções da análise gráfica são complementadas e ratificadas pela métrica de desempenho -



**Figura 07- Distribuição da previsão do retorno %**

RMSE, que foi de 0,0257 para a SVM e de 0,02221 e 0,02404 para a LSTM de menor e maior RMSE entre as 50 iterações realizadas.

A Figura 08 permite uma percepção da distribuição dos retornos por modelo. Em seguida, apesar das evidências atestadas na análise gráfica, realizamos o teste Kruskal-Wallis verificar se há igualdade entre as distribuições de retornos dos diferentes modelos e dos retornos observados. O resultado do teste Kruskal-Wallis indica a rejeição da hipótese nula, que prega a existência de igualdade, para um nível de significância de 5%. O teste de Dunn foi gerado na sequência para análise par-a-par entre as distribuições, podemos aferir que apenas as distribuições entre as iterações do modelo LSTM tiveram a hipótese nula aceita.



**Figura 08- Box-Plot da distribuição da previsão do retorno %**

Uma outra faceta para analisar os resultados num processo de tomadas de decisão de investimento, temos que caso fosse estabelecido um alvo de 3% de retorno por investimento, tem-se que das 50 iterações e que a decisão de investimento apenas seria tomada onde as situações de previsões superaram a maioria das previsões. Ou seja, consideramos para análise apenas os dias onde no mínimo 26 das 50 previsões diárias sinalizem como oportunidade de investimento, nesse escopo temos uma acurácia dos sinais de investimento de 73%. Sendo que, as decisões *naive* alcançam um *precision* de 62%, supondo que em todos dias um investidor poderia alcançar nos próximos 3 dias um retorno de 3%. Esse *target* foi calculado tomando como base a compra da ação no preço na abertura do dia seguinte e que nos próximos 3 dias haveria uma variação positiva mínima de 3%.

## 5 CONSIDERAÇÕES FINAIS

Podemos observar que o modelo proposto no presente artigo apresentou um resultado bastante interessante quando comparado aos *baselines*. Sendo os resultados considerados promissores, dado o potencial satisfatório apresentado no processo de previsão e redução de dimensionalidade de *features*. Destacamos também o desafiador e instigante campo de decisão, onde centra-se tomadas de decisões simplesmente na leitura de gráficos (imagens) com intuito de captar padrões com o objetivo de prospectar retornos futuros de um determinado ativo.

Assim, sinalizamos que as técnicas CAE e LSTM possuem um instigante e promissor potencial para auxiliar o investidor no processo de tomada de decisão. Tratam-se de técnicas que possuem uma ampla condição de exploração científica, com exigência para estudos amplos e consistentes.

É justamente com esses destaques finais, que deixamos destacado que realizaremos pesquisas futuras com o intuito de ampliar a base de dados, seja por questão temporal ou de outros ativos, refinamento dos hiperparâmetros, sistematização mais robusta dos experimentos e vislumbrar um grande potencial nas técnicas ora utilizadas.

## AGRADECIMENTOS

Este trabalho contou com o apoio do Centro Federal de Educação Tecnológica de Minas Gerais (CEFET-MG), da Universidade Federal de Lavras (UFLA) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (processo número 460.048 / 2014-7).

## REFERÊNCIAS

- CHEN, S.; HE, H. Stock prediction using convolutional neural network. In: *IOP Conference Series: Materials Science and Engineering*, v. 435, p. 1 - 9, 2018.
- DING, X.; ZHANG, Y.; LIU, T.; DUAN, J. Deep learning for event-driven stock prediction. In: *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI)*, p. 2327 – 2333, 2015.
- FAMA, E. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, v. 25, n. 2, p. 383 – 417, 1970.
- FISCHER, T.; KRAUSS, C. Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, v. 270, n. 2, p. 654 – 669, 2018.
- GAO, T.; CHAI, Y. Improving stock closing price prediction using recurrent neural network and technical indicators. *Neural Computation*, v. 30, n. 10, p. 2833 – 2854, 2018.
- GRAVES, A.; MOHAMED, A.; HINTON, G. Speech recognition with deep recurrent neural networks. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, p. 6645 – 6649, 2013.
- HOSAKA, T. Bankruptcy prediction using imaged financial ratios and convolutional neural networks. *Expert Systems with Applications*, v. 117, n. 1, p. 287 – 299, 2019.
- HU, G.; HU, Y.; YANG, K.; YU, Z.; SUNG, F.; ZHANG, Z.; XIE, F.; LIU, J.; ROBERTSON, N.; HOSPEDALES, T.; MIEMIE, Q. Deep stock representation learning: From candlestick charts to investment decisions. arXiv:1709.03803v3, p. 1 – 5, 2018.
- KIM, K-J. Financial time series forecasting using support vector machines. *Neurocomputing*, 55, p. 307-319, 2003.
- KIROS,R.; SALAKHUTDINOV, R.; ZEMEL, R. 2014. Multimodal neural language models. In: *Proceedings of the 31st International Conference on Machine Learning*, v. 32, 595–603, 2014.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural

networks. In: *Advances in Neural Information Processing Systems*, v. 25, p. 1106 – 1114, 2012.

LE, H. H.; VIVIANI, J-L. Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. *Research in International Business and Finance*, 44, p. 16–25, 2018.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, p. 436 – 444, 2015.

LV, L.; KONG, W.; QI, J.; ZHANG, J. An improved long short-term memory neural network for stock forecast. In: *2nd International Conference on Electronic Information Technology and Computer Engineering*; MATEC Web of Conferences, v. 232, p. 1 – 6, 2018.

MEDSKER, L. R.; JAIN, L. C. *Recurrent neural networks: Design and applications*. Boca Raton: CRC Press, 389, 2001.

OZTURK, H.; NAMLI, E.; ERDAL, H. I. Modelling sovereign credit ratings: The accuracy of models in a heterogeneous sample. *Economic Modelling*, 54, p. 469–478, 2016.

SAK, H.; SENIOR, A.; BEAUFAYS, F. Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. *arXiv: 1402.1128v1*, p. 1 – 5, 2014.

SHARPE, W. F. The sharpe ratio. *Journal of Portfolio Management*, v. 21, n. 1, p. 49 – 58, 1994.

SINGH, R.; SRIVASTAVA, S. Stock prediction using deep learning. *Multimedia Tools and Applications*, v. 76, n. 18, p. 18569 – 18584, 2017.

STOJANOVSKI, D.; STREZOSKI, G.; MADJAROV, G.; DIMITROVSKI, I. Twitter sentiment analysis using deep convolutional neural network. In: *Proceedings of the 10th International Conference on Hybrid Artificial Intelligent Systems (HAIS)*, p. 726 – 737, 2015.

VELAY, M. DANIEL, F. Stock chart pattern recognition with deep learning. *arXiv:1808.00418v1*, p. 1 – 6, 2018.