

Utilizando HMM para previsão de preço e estratégia de investimento em criptomoedas BitCoin

Luiz Phillip Q. Silva, André P. D. de Araújo, Daniel O. C. Cota, Gabriel O. C. Cota,
Antonio A. de A. Rocha

¹ Instituto de Computação (IC)
Universidade Federal Fluminense (UFF) – Niteroi, RJ – Brazil

{luizphillip, andrepda, danielotavio, gabrielotavio}@id.uff.br
arocha@ic.uff.br

Abstract. *The machine learning is more and more present at day a day for people. These techniques can help, for example, in decision making about investments in the stock marketing to reduce the risks associated with financial operations. New agents are more and more present in the society, the bitcoin is a clear example of technology insertion in financial transactions. But, like all coin is possible negotiate in the stock marketing focusing in its increase. In this article, three techniques of mathematical models are applied for forecasting the future Bitcoin price. Among all techniques analyzed, the Hidden Markov Model (HMM) obtained the best performance and a return on investment of more than 50,000 U\$.*

Resumo. *O aprendizado de máquina está cada vez mais presente no dia a dia da população. Essas técnicas podem auxiliar, por exemplo, na tomada de decisão sobre investimentos no mercado de ações a fim de reduzir os riscos associados às operações financeiras. Novos agentes estão cada vez mais presentes na sociedade, sendo o Bitcoin um exemplo claro da inserção tecnológica nas transações financeiras. Entretanto, como toda moeda, é possível negociar no mercado de ações, apostando em sua valorização. Neste artigo, três técnicas de modelagem matemática são utilizadas a fim de prever o valor futuro do Bitcoin. Dentre todas as técnicas analisadas o Modelo Oculto de Markov (HMM) obteve a melhor performance e um retorno de investimento de mais de 50.000 U\$.*

1. Introdução

As criptomoedas são ativos financeiros digitais que fazem uso de algoritmo e conceitos tecnológicos, tais como criptografia, monitoramento e descentralização. Essas técnicas viabilizam a realização de transações financeiras de forma confiável e sem a necessidade de uma entidade centralizadora para validar a operação. Esse conceito foi desenvolvido há pouco mais de uma década, o qual mais tarde foi responsável pelo desenvolvimento de uma nova tecnologia. No ano de 2009, sob autoria de Satoshi Nakamoto¹ [Nakamoto and Bitcoin 2008], o Bitcoin foi criado e se tornou o pioneiro das criptomoedas. Embora, desde então diversas outras soluções já tenham sido desenvolvidas

¹Satoshi Nakamoto é o codinome adotado pelo criador do Bitcoin, que até hoje tem a sua identidade real desconhecida.

(Ethereum, XRP (Ripple), Litecoin, EOS, NEO², apenas para citar algumas), o Bitcoin tem se mantido até hoje como a mais utilizada dentre todas as moedas digitais existentes.

Esses ativos virtuais têm ocupado cada vez mais espaço no mercado financeiro, através de suas transações descentralizadas, verificação simples de pagamento e segurança. Por não ser uma moeda física, todo o controle é feito através de registros armazenados em um livro-razão, baseado em uma tecnologia denominada Blockchain [Nakamoto and Bitcoin 2008]. A compra e venda de criptomoedas, em especial do *Bitcoin*, segundo [Nguyen 2017], é a forma do mercado fornecer uma moeda digital para transações econômicas entre todos os países do mundo com taxa dinâmica. Refletindo sobre a importância do uso de *Bitcoin*, [Gupta and Dhingra 2012a] afirmam que a necessidade em encontrar soluções tem levado pesquisadores a estudarem formas de superar problemas como volatilidade, sazonalidade e dependência do tempo, da economia e de outros mercados, mas principalmente o de *Bitcoin*.

Atualmente, 1 Bitcoin (ou simplesmente, BTC) vale cerca de 34.500 USD (dólares norte americanos). Porém, como todo ativo financeiro, o seu valor de mercado pode variar bruscamente devido a certos riscos. Além da volatilidade desse ativo causado pela inerência do mercado financeiro, a flutuação do preço do Bitcoin está eventualmente sujeito também a questões tecnológicas, devido a sua natureza de moeda digital. Para essa moeda, em particular, alguns problemas que podem eventualmente influenciar são, por exemplo, os riscos operacionais no caso de problemas de segurança (e.g., identificação de vulnerabilidades no protocolo) e problemas no processo de mineração (i.e., custo computacional elevado ou falta de interesse dos participantes) [Jesus et al. 2018b, Jesus et al. 2018a].

Prever os preços de alguns ativos financeiros é considerado um problema complexo devido à natureza do mercado [Ariyo et al. 2014]. Essas técnicas podem envolver tanto a previsão dos preços dos ativos, como também a tecnologia pode ser usada visando reduzir os riscos dos investimentos realizados [Ariyo et al. 2014]. Por essas características, a predição desses ativos continua sendo uma área de muito interesse dos pesquisadores. Técnicas de Inteligência Artificial e Aprendizado de Máquina - como Redes Neurais Artificiais, *Lógica Fuzzy* e Máquinas de Vetores de Suporte - têm sido amplamente utilizadas em propostas na literatura [Gupta and Dhingra 2012b]. No caso das criptomoedas, as técnicas de previsão oferecem grandes oportunidades para arbitragem³, que é uma grande força motriz para a pesquisa neste campo. Assim como no caso das ações, saber prever o valor futuro ou a tendência de aumento ou redução do valor do Bitcoin, mesmo por uma fração segundo, pode trazer grandes lucros. Ainda que essa previsão não seja probabilisticamente correta, pode ser muito lucrativa. Ao mesmo tempo, modelar e compreender o comportamento do mercado pode ainda ser de grande utilidade para identificar ocorrências anômalas na variação do preço, que levem a identificar eventuais ações maliciosas ou problemas no funcionamento do protocolo das criptomoedas.

Neste trabalho foram analisadas três técnicas de modelagem para séries temporais com o objetivo de prever o valor futuro do *Bitcoin*, foram elas: o Modelo Oculto de Markov (ou HMM, do termo em inglês *Hidden Markov Model*), o modelo autorregressivo

²<https://www.loc.gov/law/help/cryptocurrency/world-survey.php>

³Arbitragem é a realização de uma operação de compra ou venda de um ativo no mercado financeiro, com o objetivo de aumentar o lucro sobre o valor do ativo.

integrado de médias móveis (ARIMA) e uma variação da rede neural recorrente, chamado de Célula de Longa e Curta Memória (ou, em inglês, *Long Short Term Memory - LSTM*). Os resultados obtidos demonstraram que a previsão utilizando HMM pode ser um caminho factível para a previsão das variações da moeda virtual e, com isso, a redução do risco associado às operações com a criptomoeda. As contribuições científicas deste trabalho são: (i) o uso do Modelo Oculto de Markov para prever o preço do Bitcoin de forma inédita, utilizando o Log Retornos como variável observada no modelo HMM; (ii) uma análise comparativa do desempenho do método proposto usando HMM junto a diversas técnicas utilizadas para prever o comportamento de séries temporais. Para isso, foram utilizadas métricas como raiz do erro médio quadrático (RMSE), erro médio quadrático logarítmico (MSLE), erro médio absoluto percentual (MAPE) e Coeficiente U de Theil; (iii) Por fim, propõe-se uma estratégia ingênua de investimento, com o HMM proposto, e comparando-o com adaptações dos outros mecanismos de previsão, para avaliar o retorno do investimento.

O restante deste trabalho está dividido da seguinte forma: a Seção 2 corresponde à fundamentação teórica, em que são descritos os modelos de previsão utilizados. A Seção 3 apresenta os trabalhos relacionados. Nas Seções 4 e 5 são apresentados, respectivamente, as propostas para um modelo de previsão de preços de Bitcoin utilizando HMM e para um mecanismo ingênuo de investimento. A Seção 6 consiste na avaliação dos métodos propostos. Finalmente, as conclusões e proposta de trabalhos futuros são apresentados na Seção 7.

2. Modelos de previsão

Neste trabalho será apresentado uma comparação entre diversas técnicas de previsão existentes na literatura. O intuito desta comparação é demonstrar a eficiência dos Modelos Ocultos de Markov para esta tarefa. Foram utilizados 3 modelos distintos, porém em dois casos variações da modelagem foram utilizadas pois são comumente referenciadas na literatura.

O primeiro modelo analisado foi o Autorregressivo, Integrado de Média Móvel (ARIMA). Este é um dos modelos mais simples existentes na literatura para realizar previsões. Para tal modelagem é necessário determinar os hiper parâmetros, os quais são escolhidos pela autocorrelação, autocorrelação parcial e a quantidade de diferenças necessárias para tornar a série temporal estacionária [Wu et al. 2018]. Modelo Linear é composto de três componentes, autorregressivo (AR), média móvel (MA) e integrado (D).

$$AR(p) := \sum_{i=1}^p \alpha_i x_{t-i} + \epsilon(t) \quad (1)$$

$$MA(q) := \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon(t) \quad (2)$$

Sendo que, o termo integrado consiste de uma diferenciação na série para que a série temporal seja estacionária;

$$ARIMA(p, d, q) = AR(p) + MA(q) \quad (3)$$

A segunda abordagem utilizada foi a Célula de Longa e Curta Memória (*LSTM*). Esse tipo de abordagem representa o estado da arte do aprendizado sequencial [Wu et al. 2018]. A principal característica desse tipo de modelo é compreender o quanto um passado mais distante é capaz de afetar os valores futuros da criptomoeda analisada. Para esta abordagem é necessário também definir alguns hiper parâmetros. Porém, diferente do *ARIMA*, não há um teste estatístico que possa informar quais parâmetros devem ser utilizados. Em geral, esses hiper parâmetros são selecionados pela experiência do pesquisador na área, ou por tentativa e erro.

A terceira, e última, abordagem investigada neste artigo foi o uso de HMM para a previsão do preço do Bitcoin. Segundo [Nguyen 2017], o *Hidden Markov Model* é um modelo de detecção de sinal que foi introduzido em 1966 por [Baum and Petrie 1966]. Nos Modelos Ocultos de Markov, os dados da sequência são comumente chamados de estado observado. Este estado é então transformado, através de funções de probabilidade, em dados não observados (estado oculto). Esse estado oculto pode ser referenciado como uma técnica de agrupamento em séries temporais.

Os elementos básicos de um HMM para múltiplas observações são:

- Dados de observação: $O = \{O_t^{(l)}, t = 1, 2, \dots, T, l = 1, 2, \dots, L\}$, onde l é um número de independentes sequências de observação e T é o comprimento de cada sequência;
- Sequência de estado oculto de O : $Q = \{q_t, t = 1, 2, \dots, T\}$;
- Valores possíveis de cada estado S_i : $\{S_i, i = 1, 2, \dots, N\}$;
- Possíveis símbolos por estado: $\{v_k, k = 1, 2, \dots, M\}$;
- Matriz de transição: $A = \{(a_{ij})\}$, onde $a_{ij} = P(q_t = S_j | q_{t-1} = S_i) \forall i, j = 1, 2, \dots, N$;
- Probabilidade inicial de estar no estado (regime): $\{S_i \rightarrow t = 1, p = p_i\}$, onde $p_i = P(q_1 = S_i), i = \{1, \dots, N\}$;
- Matriz de probabilidade de observação: $B = b_i(k)$, onde $b_i(k) \equiv b_i\{(O_t = v_k) \equiv P(O_t = v_k | q_t = S_i)\}, i = \{1, \dots, N\}, k = \{1, \dots, M\}$.

Os parâmetros de um HMM são as matrizes A e B e o vetor p . Por conveniência, utiliza-se uma notação compacta para os parâmetros, dada por:

$$\pi \equiv \{A, B, p\}. \quad (4)$$

Se a transformação para o estado oculto assume uma distribuição gaussiana, então temos um HMM contínuo com $b_i(k) = b_i(O_t = v_k) = \mathcal{N}(v_k, \mu_i, \sigma_i^2)$, onde \mathcal{N} é a função de densidade gaussiana com média (μ_i) e a variância (σ_i^2) da distribuição correspondente ao estado S_i . Por conveniência, é escrito $b_i(O_t = v_k)$ como $b_i(O_t)$. Então, os parâmetros do HMM são: $\pi = \{A, \mu, \sigma^2, p\}$. Onde, μ e σ^2 são vetores de médias e variâncias das distribuições gaussianas, respectivamente. Com a suposição de que as observações são independentes, a probabilidade de observação, denotada por $P(O|\lambda)$, é:

$$P(O|\lambda) = \prod_{l=1}^L P(O^{(l)}|\lambda) \quad (5)$$

Para ajustar os parâmetros do modelo acima descrito dois algoritmos são utilizados: *Forward* e *Backward*. Essas técnicas determinam A , μ e σ^2 de forma que cada

observação possua um estado bem definido. Os dois algoritmos para ajuste de parâmetros podem ser encontrados em [Nguyen 2017].

3. Trabalhos Relacionados

Na literatura, encontram-se aplicações do Modelo Oculto de Markov para previsão de preços de diversos ativos. Os autores de [de Souza e Silva et al. 2010] apresentam como o HMM pode ser utilizado para retornos eficientes de investimentos no mercado de petróleo. Para tal, utilizaram transformações de Wavelet na série temporal para um melhor ajuste do HMM. Essa técnica possibilitou ganhos superiores a um investimento ingênuo, mesmo em épocas de queda do mercado. Esse tratamento nos dados antes da modelagem também é observada em [Nguyen 2017]. Os autores neste trabalho mostram os benefícios de se utilizar o cálculo de retornos diários invés da utilização da série financeira nas hipóteses assumidas quando se utiliza HMM. Neste caso, o modelo mostrou-se extremamente eficiente e com resultados de retornos muito favoráveis em empresas de tecnologia.

Os trabalhos envolvendo a previsão do preço de criptomoedas têm crescido ao longo dos anos. Em 2018, [Wu et al. 2018] propuseram a elaboração de uma nova estrutura de modelo LSTM para a previsão diária do preço do Bitcoin. No mesmo ano, [Chappell et al. 2019] utilizou o modelo Markov-Switching para identificação da dinâmica de volatilidade das séries temporais do preço do Bitcoin. Nesta mesma linha, [Maupin 2019] publicou um trabalho em 2019, envolvendo a implementação da abordagem Markov-Switching para prever o preço de quatro criptomoedas: Bitcoin, Ethereum, Ripple e Litecoin. Neste mesmo ano, [Koki et al. 2019] propuseram realizar a previsão das criptomoedas Bitcoin e Ether através de um modelo oculto de markov não-homogêneo, ou seja, modelo no qual a matriz de transição estados do HMM não é constante ao longo do tempo. No ano seguinte, [Koki et al. 2020], publicaram outro artigo na qual propuseram uma previsão dos valores do Bitcoin, Ethereum e Ripple com vários modelos ocultos de markov diferentes.

Dentre os artigos que comparam diferentes algoritmos para modelar a previsão de criptomoedas, destacam-se: [Hua 2020] que propôs a previsão de preço do Bitcoin utilizando dois modelos distintos que são o ARIMA e o LSTM, sendo o LSTM com melhor *performance* principalmente em previsões de instantes de tempo mais longa. O autor em [Xu 2020] utilizou LSTM, GRU e HMM para prever o preço do Bitcoin. Ao final, a comparação dos resultados demonstra que o LSTM possui um RMSE 81% maior que o HMM, o qual foi o melhor modelo. O HMM proposto por [Xu 2020] foi reproduzido nas próximas seções afim de ter seu desempenho comparado com outra estratégia de HMM implementada neste artigo.

A estrutura do HMM é bastante versátil e vem sendo utilizada em diversas áreas do conhecimento. Trabalhos como [Ghosh et al. 2019] demonstram como a capacidade preditiva dessa classe de modelos pode ser utilizada para manutenção preditiva, utilizando uma técnica conhecida como *Digital Twin*. Essa linha de raciocínio também é seguida por [Li et al. 2020], no qual são utilizadas as informações geradas pelo modelo para prever o tempo de vida restante do equipamento. Vale ressaltar que, no passado, essa classe de modelos era muito utilizada para o processamento de falas e sinais, como apresentado em [Varga and Moore 1990, Schuller et al. 2003, Van Compernelle 1989], ou para previsão de estatísticas de perda aplicados a mecanismos de correção de erros de pacotes na

Internet [Silveira and de Souza e Silva 2012].

4. Proposta de Mecanismos de Previsão

Esta seção apresenta a utilização do HMM para prever os preços do *Bitcoin*. Essa previsão foi realizada usando duas estratégias de HMM. A primeira é a reprodução, a título de comparação, da proposta apresentada por [Xu 2020] e a segunda técnica é a contribuição deste artigo, que toma como inspiração o que é feito no mercado de ações por [Nguyen 2017], porém adaptado para o objetivo de previsão do preço de *Bitcoin*. A avaliação do desempenho da previsão é medida comparando a resposta do modelo com o comportamento real do mercado, determinado pelo cálculo dos erros relativos.

O modelo proposto por [Xu 2020] recebe como variáveis observadas as relações entre os valores de fechamento, mais alto, mais baixo e com o preço de abertura. A Equação (6) apresenta como são obtidas as variáveis observadas neste caso.

$$O_t = \left(\frac{c_t - a_t}{a_t}, \frac{h_t - a_t}{a_t}, \frac{a_t - b_t}{a_t} \right) \quad (6)$$

onde c_t, a_t, h_t e b_t representam o preço de fechamento, abertura, mais alto e mais baixo do dia t .

Já no modelo proposto neste artigo, baseado na ideia de retornos apresentada por [Nguyen 2017], as variáveis dependem do retorno de cada uma das medidas (preço de fechamento, abertura, mais alto e mais baixo do dia). Ou seja, o quanto os preços variaram de um dia para o outro. Em nossa proposta, o Log Retornos é utilizado para mensurar essa variação e é calculado da seguinte forma:

$$\text{LogRetornos} = \ln \left(\frac{X_t}{X_{t-1}} \right). \quad (7)$$

Onde, X_t e X_{t-1} são os valores de cada uma das 4 variáveis consideradas, no dia t e no dia $t - 1$, respectivamente.

Com o cálculo dos Log Retornos executado, basta modelar o processo com base nas novas medidas observadas. A Equação (8) expõe como é feita a modelagem dos dados observados.

$$O_t = (a_t^{\text{ret}}, h_t^{\text{ret}}, b_t^{\text{ret}}, c_t^{\text{ret}}) \quad (8)$$

Onde, $c_t^{\text{ret}}, a_t^{\text{ret}}, h_t^{\text{ret}}$ e b_t^{ret} representam os Log Retornos do preço de fechamento, abertura, mais alto e mais baixo do dia t , respectivamente.

O modelo de HMM não possui a capacidade de realizar previsões, mas sim de encontrar observações similares. Por isso, para cada modelagem será necessário utilizar uma técnica específica para realizar a previsão da criptomoeda.

O modelo proposto por [Xu 2020] utiliza a seguinte técnica para a previsão:

1. Ajuste dos parâmetros do HMM e calcule a probabilidade do modelo;
2. Encontre o dia ($t - 1$) no passado que possua a máxima verossimilhança (mínima logaritmo da verossimilhança) com o dia atual (t);
3. A variação que acontecerá entre t e $t + 1$ será a mesma que ocorreu entre $t - 1$ e t ;

Já o nosso modelo, no caso da utilização do Log Retornos, uma abordagem mais simples é utilizada para realizar as previsões, sendo os três passos são modificados para:

1. Ajuste dos parâmetros do HMM e calcule a probabilidade do modelo;
2. Encontre o dia ($t - 1$) no passado que possua a máxima verossimilhança (mínima logaritmo da verossimilhança) com o dia atual (t);
3. O log retornos de t será o mesmo que ocorreu em $t - 1$;

5. Método Ingênuo de Retorno

Investir em ativos financeiros, em geral, está relacionada a ganhos monetários. Os métodos propostos de previsão podem auxiliar na tomada de decisão do melhor momento de se investir num determinado ativo ou conjunto de ativos financeiros. Em geral são cobradas taxas por cada operação de compra e venda, além de outros pagamentos de tributos ao governo. Por essa complexidade adicional, a qual depende do tipo de investimento realizado e do quanto está sendo investido, um método ingênuo de retorno de investimento é proposto como forma de comparar cada tipo de algoritmo num mesmo período de tempo. Neste método não serão consideradas as taxas pagas pelo investidor tanto para a corretora, quanto para o governo federal. Porém, eles podem ser facilmente adaptados para levarem em consideração essas taxas.

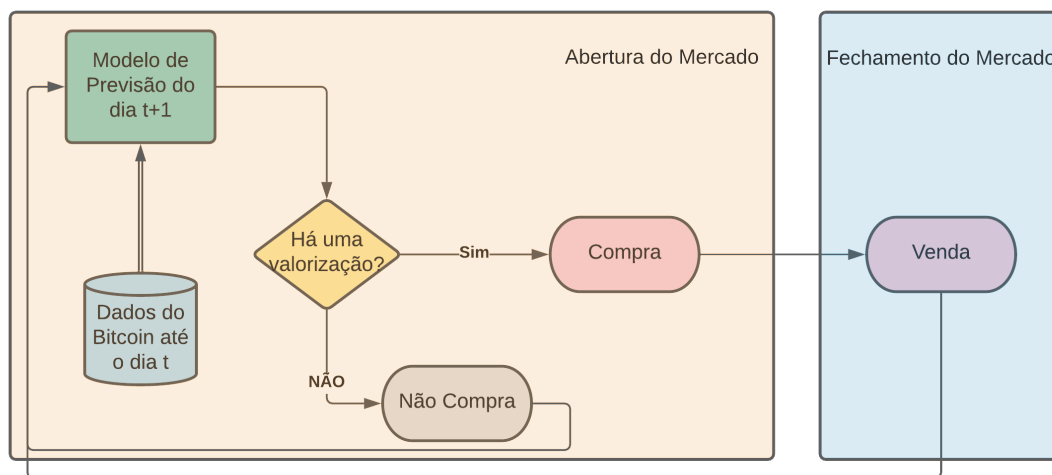


Figura 1. Método do Retorno do Investimento

O método ingênuo de retorno define a operação de decisão de compra/venda do Bitcoin a partir do modelo de previsão. A Figura 1 ilustra um diagrama de fluxo do método do retorno do investimento proposto, que pode ser descrito da seguinte forma. A cada momento t será realizada uma previsão para o dia $t + 1$ (utilizando o Modelo de Previsão): (i) Caso o valor de fechamento previsto do dia $t + 1$ seja maior que o preço de fechamento do dia atual(t), então compra-se a criptomoeda com o valor de abertura do dia $t + 1$ (que é exatamente igual ao valor de fechamento do dia t). Após todo o dia de flutuação da moeda, tudo que foi comprado na abertura do mercado será vendido na hora de fechamento do mercado pelo valor de fechamento do dia $t + 1$; (ii) Caso o preço no dia $t + 1$ seja menor que no dia t , nada será feito até o final do dia $t + 1$, quando

novamente se avaliará a decisão de compra ou venda para o dia seguinte com base no valor de fechamento. Note que, se a decisão de compra será mantida (ou seja, o investidor decidiu por comprar no dia t e o modelo prever que haverá uma valorização novamente no dia $t + 1$), na prática, não precisaríamos realizar a venda no fechamento do dia t para comprar novamente na abertura do dia $t + 1$. Mas, por uma questão didática, mentemos essa lógica. Ao final de cada dia, serão acrescentados dados à base de dados do Bitcoin, que alimenta o modelo, e com isso uma previsão para o dia seguinte será realizada já com os dados atualizados.

6. Validação do Modelo Previsão e Método de Investimento

Neste trabalho foram utilizadas 5 métricas para a validação modelo previsão de preço do Bitcoin e do mecanismo ingênuo de investimento. Além do modelo proposto utilizando HMM com Log Retornos, outros modelo de previsão também foram avaliados, a título de comparação. Também à título de comparação, os mesmos modelos foram considerados como fonte de decisão para um mecanismo ingênuo de investimento.

6.1. Métricas

As métricas consideradas foram: erro médio quadrático (RMSE), erro médio quadrático logarítmico (MSLE), absoluto percentual (MAPE), Coeficiente U de Theil e Retorno do investimento (RI). Todas as métricas foram analisadas no conjunto de teste. Ou seja, um conjunto para o qual o modelo não teve acesso antes de realizar a previsão.

A métrica mais simples analisada é a RMSE. Essa forma de validação representa o quão distante o modelo está da realidade, penalizando grandes desvios. A Equação (9) apresenta a fórmula do RMSE.

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^n (real_i - predito_i)^2} \quad (9)$$

Já o MSLE trata-se de uma medida muito utilizada nos casos em que há um aumento substancial nos valores a serem preditos. É importante ressaltar que subestimação será mais penalizadas que superestimação, diferente do que ocorre com o RMSE. A Equação (10) apresenta a fórmula desta métrica.

$$MSLE = \frac{1}{N} \cdot \sum_{i=1}^n (\log(1 + real_i) - \log(1 + predito_i))^2 \quad (10)$$

O MAPE, também utilizado, é uma medida que fornece uma capacidade de interpretação dos resultados mais eficiente, tendo em vista que é possível analisar os resultados independentemente da ordem de grandeza dos valores reais e preditos. A fórmula do MAPE é apresentada na Equação (11).

$$MAPE = \frac{1}{N} \cdot \sum_{i=1}^n \left(\frac{real_i - predito_i}{real_i} \right) \cdot 100 \quad (11)$$

A Equação (12) apresenta o Coeficiente U de Theil. A ideia desta métrica de validação é comparar o quanto um modelo em questão é melhor ou pior do que um previsor ingênuo. Esse previsor ingênuo é simplesmente a replicação do valor anterior. Quando

essa métrica está acima de 1 significa que, de forma geral, o modelo desenvolvido possui uma *performance* pior que o previsor ingênuo.

$$UTheil = \sqrt{\frac{\sum_{i=1}^N (real_i - pred_i)^2}{\sum_{i=1}^N (real_i - real_{i-1})^2}} \quad (12)$$

A última métrica utilizada foi o Retorno do investimento (RI). Essa métrica foi considerada especificamente para avaliar o desempenho dos mecanismos ingênuos de investimento. Ela é desenvolvida a partir do somatório do lucro (prejuízo) acumulado ao longo do tempo em que a decisão de compra

$$RI = \sum_i (VV_i - VC_i) \quad (13)$$

Onde, VC_i e VV_i são, respectivamente, os valores de compra e venda do Bitcoin no dia i de compra da criptomoeda informado pelo modelo.

Cada uma dessas métricas mede uma característica relevante dos modelos. No caso do MAPE e RMSE essas medidas são capazes de informar o quanto em média o modelo falha em prever o valor correto independente da característica do problema. Já o MSLE é utilizada em problemas com alta volatilidade, desta maneira é possível entender a capacidade de adaptação do modelo. O Coeficiente U de Theil demonstra o quão eficiente é o modelo em comparação com o último valor da série temporal. Por fim, o retorno do investimento tem sua aplicação neste trabalho por se tratar de uma série financeira, sendo o retorno do capital um objetivo comum nesta área. As seções a seguir apresentam os resultados obtidos para cada modelo.

6.2. Resultados

Os valores diários do Bitcoin (abertura, baixa, alta, fechamento) foram obtidos através de uma API⁴. A partir do histórico diário da criptomoeda foram aplicados os modelos descritos na Seção 5. A Figura 2 expõe o comportamento histórico da criptomoeda entre fevereiro de 2016 e fevereiro de 2021, apresentando ao longo dos dias o preço da moeda e o volume de negociações. Como pode-se observar houve um aumento na quantidade de negociações ao longo do tempo. Além disso, após a forte desvalorização entre 2017-2018 há uma tendência de alta em seus preços. Outra informação relevante apresentada na Figura 2 é a separação entre os dados utilizados para treinamento dos modelos e para teste. O comportamento da criptomoeda entre fevereiro de 2016 e fevereiro de 2020 foi utilizado para treino, enquanto que os dados posteriores ao último dia da base de treino até fevereiro de 2021 foi utilizado para teste dos modelos.

O primeiro modelo treinado foi o ARIMA, o qual foi escolhido por ser um modelo linear e que seus hiperparâmetros são facilmente determinados a partir de análises estatísticas. Este modelo possui três parâmetros a serem definidos pelo usuário e, comumente, o modelo é descrito da forma ARIMA(p,q,i), em que p corresponde a quantidade de termos usado no modelo auto regressivo, q corresponde a quantidade de termos utilizados no modelo de média móvel e i representa se é necessário diferenciar a série temporal

⁴<http://messari.io/>



Figura 2. Visualização Histórica do Comportamento da Cripto Moeda

para torna-lá estacionária. Os testes de autocorrelação, autocorrelação parcial e Dickey Fuller são capazes de fornecer os parâmetros necessários para a modelagem.

A *performance* do primeiro modelo, ARIMA Sem Retreino, com os parâmetros selecionados foi baixa, atingindo um *MAPE* de $\approx 27\%$. Portanto, uma abordagem utilizando um retreinamento foi utilizada. O modelo é obtido da seguinte maneira:

1. O modelo é treinado com os dados de treino e parâmetros selecionados para o ARIMA Sem Retreino;
2. Faz-se a previsão do primeiro valor do conjunto de teste;
3. Após a previsão, o valor real é unido aos dados de treino e o preço mais antigo do Bitcoin é retirado do conjunto;
4. O modelo então é retreinado e faz-se uma nova previsão e retorna ao passo anterior;

Essa rotina é utilizada até que a última previsão do conjunto de teste seja executada. Com essa rotina o modelo ARIMA Com Retreino obteve resultados significativamente melhores, com um *MAPE* de 2.6%. Vale ressaltar que esse tipo de abordagem pode ser utilizada e as métricas são válidas já que o modelo só recebe uma atualização após realizar uma previsão.

O segundo modelo implementado foi o LSTM. Diferente do ARIMA em que se há testes estatísticos capazes de determinar os parâmetros do modelo, para o aprendizado profundo é necessário realizar testes com diferentes arquiteturas até que se encontre a melhor. Após a determinação da melhor arquitetura o modelo foi testado com o conjunto de teste e seus resultados foram similares ao do ARIMA Com Retreino, obtendo um *MAPE* de 2.63

A última abordagem analisada foi o HMM. Para este tipo de modelo duas formas de representar as variáveis observadas foram utilizadas, uma baseada em [Xu 2020] que será descrita por "HMM Yike_Xu" nos gráficos apresentados e outra que foi utilizada em [Nguyen 2017] que será apresentado como "HMM Proposto". A principal diferença

entre as duas abordagens é que a primeira considera apresentar as variáveis observadas em função do preço de abertura, enquanto que a segunda utiliza a variação ocorrida entre dois dias seguidos como variável observada. Outra diferença crucial é na forma como o preço futuro é predito, enquanto que na abordagem proposta por [Nguyen 2017] a própria variação já forneceria esta informação, para o modelo de [Xu 2020] é necessário encontrar o dia no passado mais parecido com o dia atual e só então calcular a variação para obter o preço futuro.

Em termos de *performance* o “HMM Yike_Xu” obteve um *MAPE* mais próximo, porém menor, que os outros dois modelos apresentados. Isso já mostra que mesmo que a representação das variáveis observadas não auxilie o modelo na predição, a abordagem já fornece resultados consistentes. Em relação ao “HMM Proposto” o *MAPE* obtido foi 42% menor que o mesmo indicador do ARIMA Com Retreino e do LSTM e 38% menor que do “HMM Yike_Xu”. Essa característica reforça a necessidade de uma boa representação das variáveis observadas a fim de obter resultados consistentes e superiores a outros modelos quando de utiliza o HMM.

Os valores reais e preditos por cada modelo no conjunto de teste são apresentados na Figura 3. Sobre esse gráfico alguns pontos valem ser destacados:

- O modelo ARIMA Sem Retreino possui uma previsão praticamente constante;
- O período (Fevereiro/2020-Fevereiro/2021) foi selecionado pela alta variação da cripto moeda;
- Exceto o ARIMA Sem Retreino, todos os modelos possuem uma boa capacidade preditiva, capazes de acompanhar a tendência do Bitcoin sem erros drásticos;

Desta forma, dependendo da aplicação ao qual os modelos estarão sujeitos uma abordagem mais simples, como ARIMA, pode ser utilizada.

Outra *performance* analisada graficamente é o Retorno do Investimento Acumulado. A Figura 4 apresenta o comportamento temporal desta métrica para cada modelo investigado. Um fato interessante sobre essa análise é que o pior modelo em relação ao *MAPE*, ARIMA Sem Retreino, torna-se o segundo melhor modelo para investimento. Isso pode ser justificado pela forte tendência de alta da cripto moeda analisada. Pois, se a moeda está na maior parte do tempo aumentando seu valor e o modelo sempre informa que uma compra deve ser feita, então o retorno do investimento tenderá a ser elevado. A Figura 4 ainda apresenta algumas características dos modelos que valem ser destacadas, como por exemplo:

- O modelo “HMM Proposto” obteve rendimentos superiores a todos os outros modelos a partir do dia 24 de março de 2020;
- Os modelos ARIMA Com Retreino e LSTM permaneceram a maior parte do tempo com retornos negativos (perdas);
- O modelo “HMM Yike_Xu” acompanhou na maior parte do tempo o ARIMA Sem Retreino;

As características obtidas com esses modelos ressalta a importância de investigações particulares para cada problema. Tendo em vista que, nesse caso, uma série temporal com forte tendência de alta foi utilizada para comparar os modelos.

Por fim, os resultados das 5 métricas estão resumidas na Tabela 1. Como pode ser observado o “HMM Proposto” obteve as melhores métricas 4 vezes, sendo equiparável

aos outros modelos no MSLE. O segundo melhor modelo é então o “HMM Yike_Xu”, obtendo a melhor métrica, retirando o ”HMM Proposto”, 3 vezes. Vale ressaltar que somente os Modelos Ocultos de Markov obtiveram o Coeficiente U de Theil menor que 1, ou seja, por essa métrica somente esses modelos seriam melhores que um predictor ingênuo.

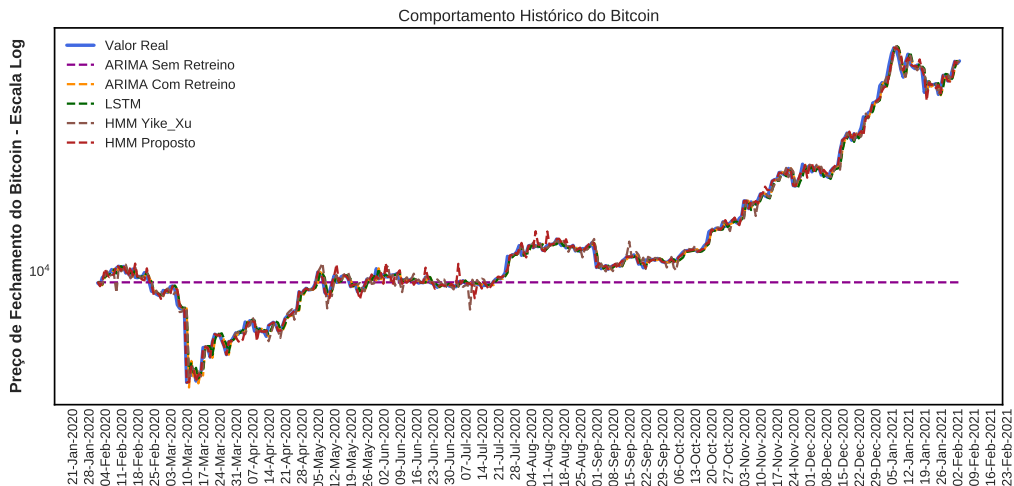


Figura 3. Comportamento Temporal do Preço de Fechamento e das Previsões de cada Modelo Analisado

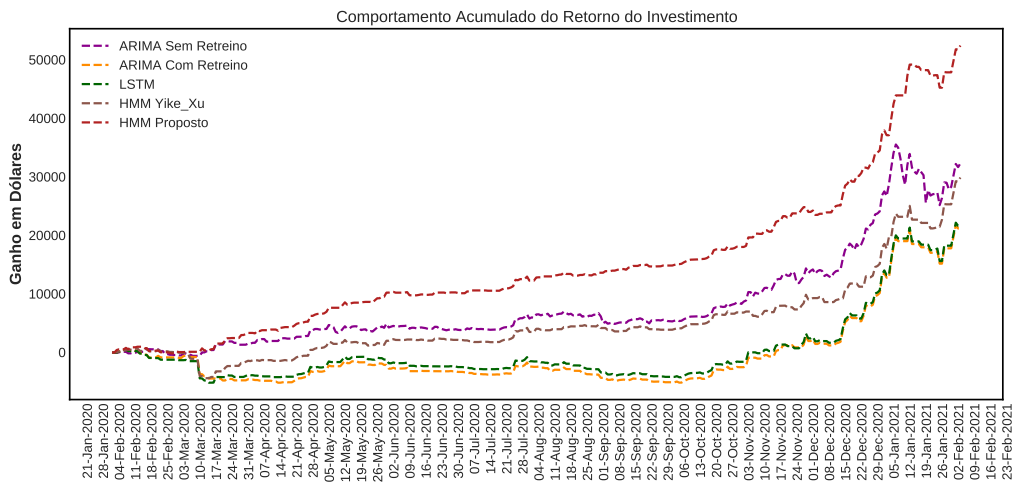


Figura 4. Retorno do Investimento Acumulado

Tabela 1. Resultados dos Modelos Analisados

| Métrica | ARIMA s/ Retreino | ARIMA c/ Retreino | LSTM | HMM Proposto | HMM Yike Xu |
|---------|-------------------|-------------------|-----------|--------------|-------------|
| U Theil | 13,36 | 1,01 | 1,03 | 0,80 | 0,91 |
| RMSE | 9.322,35 | 705,19 | 721,76 | 558,39 | 638,73 |
| MSLE | 0,0018 | 0,0018 | 0,0018 | 0,0008 | 0,0018 |
| MAPE | 26,98 | 2,60 | 2,63 | 1,50 | 2,41 |
| RI | 32.271,58 | 21.144,71 | 21.595,43 | 52.402,16 | 29.851,62 |

7. Conclusões

Neste trabalho foram apresentadas três técnicas de modelagem para séries temporais, sendo que em duas técnicas variantes foram aplicadas a fim de melhorar a *performance* dos modelos. Ao final da determinação dos hiper parâmetros necessários para cada modelo, foram analisadas os resultados através das 5 métricas propostas nesse trabalho. Os indicadores obtidos mostram que o HMM com variáveis observadas igual ao Log Retornos dos preços de abertura, fechamento, mais alto e mais baixo é o melhor modelo para o problema em questão.

As contribuições científicas deste trabalho são a utilização de Modelos Ocultos de Markov para prever o preço de fechamento do Bitcoin, utilização do Log Retornos como variável observada no modelo HMM, implementação de métricas como Coeficiente U de Theil e Retorno do Investimento para validação dos modelos e um comparativo com diversas técnicas utilizadas para prever o comportamento de séries temporais. Vale destacar que o HMM com Log Retornos obteve um MAPE de 1.50% e um retorno de \approx \$ 52.400, 00 para um investimento inicial de \approx \$ 10.000, 00 durante 1 ano.

Trabalhos futuros podem utilizar deste mesmo modelo, porém com variáveis observadas distintas. Outra linha em que o HMM pode ser aplicado é no desenvolvimento de portfólios financeiros. Estratégias mais inteligentes sobre a decisão de investimento também podem ser analisadas, utilizando modelos de classificação ou aprendizado por reforço. A aplicação nos dados intra-diários também é uma possibilidade, a qual pode ser unida com a análise de sentimento de notícias do mercado financeiro.

Referências

- Ariyo, A. A., Adewumi, A. O., and Ayo, C. K. (2014). Stock price prediction using the arima model. In *2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, pages 106–112. IEEE.
- Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state markov chains. *The annals of mathematical statistics*, 37(6):1554–1563.
- Chappell, D. R. et al. (2019). Regime heteroskedasticity in bitcoin: A comparison of markov switching models. *Journal of Applied Economic Sciences (JAES)*, 14(65):621–641.
- de Souza e Silva, E. G., Legey, L. F., and de Souza e Silva, E. A. (2010). Forecasting oil price trends using wavelets and hidden markov models. *Energy Economics*, 32(6):1507–1519.
- Ghosh, A. K., Ullah, A. S., and Kubo, A. (2019). Hidden markov model-based digital twin construction for futuristic manufacturing systems. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 33(3):317–331.
- Gupta, A. and Dhingra, B. (2012a). Stock market prediction using hidden markov models. In *2012 Students Conference on Engineering and Systems*, pages 1–4. IEEE.
- Gupta, A. and Dhingra, B. (2012b). Stock market prediction using hidden markov models. *StockStock*, pages 1–4.
- Hua, Y. (2020). Bitcoin price prediction using arima and lstm. In *E3S Web of Conferences*, volume 218. EDP Sciences.

- Jesus, E., Chicarino, V., Albuquerque, C., and Rocha, A. (2018a). A survey of how to use blockchain to secure internet of things and the stalker attack. *Security and Communication Networks*, 2018:1–27.
- Jesus, E. F., Chicarino, V. R. L., Albuquerque, C., and de A. Rocha, A. A. (2018b). Stalker - uma nova estratégia para o atacante egoísta em blockchains. In *Anais do XXXVI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, Porto Alegre, RS, Brasil. SBC.
- Koki, C., Leonardos, S., and Piliouras, G. (2019). Do cryptocurrency prices camouflage latent economic effects? a bayesian hidden markov approach. In *Multidisciplinary Digital Publishing Institute Proceedings*, volume 28, page 5.
- Koki, C., Leonardos, S., and Piliouras, G. (2020). Exploring the predictability of cryptocurrencies via bayesian hidden markov models. *arXiv preprint arXiv:2011.03741*.
- Li, R., Arzaghi, E., Abbassi, R., Chen, D., Li, C., Li, H., and Xu, B. (2020). Dynamic maintenance planning of a hydro-turbine in operational life cycle. *Reliability Engineering System Safety*, 204:107129.
- Maupin, T. (2019). Can bitcoin, and other cryptocurrencies, be modeled effectively with a markov-switching approach?
- Nakamoto, S. and Bitcoin, A. (2008). A peer-to-peer electronic cash system. *Bitcoin*.— URL: <https://bitcoin.org/bitcoin.pdf>, 4.
- Nguyen, N. (2017). An analysis and implementation of the hidden markov model to technology stock prediction. *Risks*, 5(4):62.
- Schuller, B., Rigoll, G., and Lang, M. (2003). Hidden markov model-based speech emotion recognition. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)*, volume 2, pages II–1.
- Silveira, F. and de Souza e Silva, E. (2012). Predicting packet loss statistics with hidden markov models for fec control. *Computer Networks*, 56(2):628–641.
- Van Compernelle, D. (1989). Noise adaptation in a hidden markov model speech recognition system. *Computer Speech Language*, 3(2):151–167.
- Varga, A. P. and Moore, R. K. (1990). Hidden markov model decomposition of speech and noise. In *International Conference on Acoustics, Speech, and Signal Processing*, pages 845–848 vol.2.
- Wu, C., Lu, C., Ma, Y., and Lu, R. (2018). A new forecasting framework for bitcoin price with lstm. In *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 168–175.
- Wu, C.-H., Lu, C.-C., Ma, Y.-F., and Lu, R.-S. (2018). A new forecasting framework for bitcoin price with lstm. In *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 168–175. IEEE.
- Xu, Y. (2020). *Bitcoin Price Forecast Using LSTM and GRU Recurrent networks, and Hidden Markov Model*. PhD thesis, UCLA.