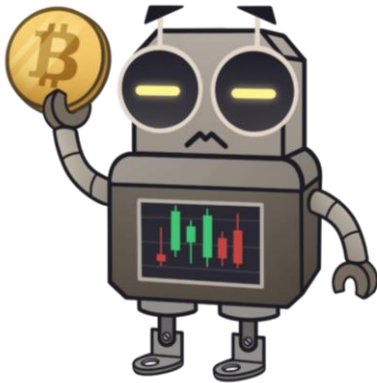


Design Robô	
Nome do Robô	Satobot
Explicação Nome	O nome faz referência ao criador do Bitcoin, Satoshi Nakamoto. Além disso, o design do robô foi inspirado na possível aparência de Satoshi Nakamoto.
Explicação Lógica Estratégia	Uma estratégia de trading a partir de um modelo binário de classificação que consiste na compra ou venda do ativo com base nos valores obtidos para quatro indicadores de lucratividade e sentimento dos investidores. A ideia é de que, por meio desses indicadores, possamos prever se o mercado está sobreavaliado ou subavaliado, ou ainda se o mercado está em euforia ou desespero, e obter retornos com isso.
Tipo de Estratégia	Trade
Classe de Ativos	Criptoativos
Universo	Bitcoin
Média Trades por mês	Máximo de 30 trades por mês
Holding Period	Mínimo um dia
Qual Plataforma Testou a estratégia	Python
Benchmark Estratégia:	Buy-and-Hold do Bitcoin

1. Resumo da Estratégia

A estratégia é um algoritmo de trading para operar Bitcoin, no qual a tomada de decisão é feita com base em um modelo de Árvore de Decisão, que é alimentada por indicadores de lucratividade e sentimento selecionados a partir de consultas à literatura disponível, sendo eles: *Spent Output Profit Ratio*, *Net Unrealized Profit-Loss*, *Coin Days Destroyed* e o número de pesquisas da palavra “Bitcoin” no *Google Trends*. Tais indicadores incorporam informações relacionadas ao universo Bitcoin, logo, a partir destes o modelo irá indicar o *trade-off* entre momentos de compra ou venda.

2. Estrutura do Modelo

2.1. Universo de ativos analisados

O universo do modelo é o Bitcoin, que foi a primeira criptomoeda criada em meados de 2009. Dentre os objetivos desse ativo digital podemos citar: funcionar como uma moeda descentralizada, sem intermediação de uma instituição financeira, como um banco; sem limitações de transações entre locais distintos e barateamento de custos. O criptoativo funciona através da tecnologia “*Blockchain*” - sistema de banco de dados que permite que os dados sejam transmitidos entre todos os participantes da rede de maneira transparente e descentralizada. Dentre as motivações para escolher o Bitcoin como ativo a ser operado, está o crescente interesse de instituições monetárias e não monetárias no mesmo em função de uma demanda cada vez mais latente pelo criptoativo, e o fato de ser a criptomoeda com maior quantidade de dados.

2.2. Variáveis do modelo

Após pesquisar e estudar diversas estratégias e indicadores do mercado de criptomoedas em artigos científicos, sites especializados, etc. Selecionamos um grupo de indicadores que acreditamos que possuem uma boa base teórica e lógica para prever variações futuras do Bitcoin, incluindo indicadores de sentimento e de lucros realizados, que são duas categorias bastante estudadas e conhecidas de indicadores de criptomoedas. Diversos artigos como *Deep neural networks for cryptocurrencies price prediction* (Spilak, 2018) testaram e chegaram a bons resultados, o que nos inspirou a desenvolver o modelo.

Google Trends

Uma das variáveis que alimenta o modelo é o número de pesquisas do termo “Bitcoin” no Google. Esse dado é obtido através do *Google Trends*, uma ferramenta que permite verificar as tendências de pesquisa no site de busca do Google com base em palavras-chave. Assim, foi identificada uma relação entre essas buscas e o comportamento da série histórica do ativo, conforme a figura (1) a seguir.

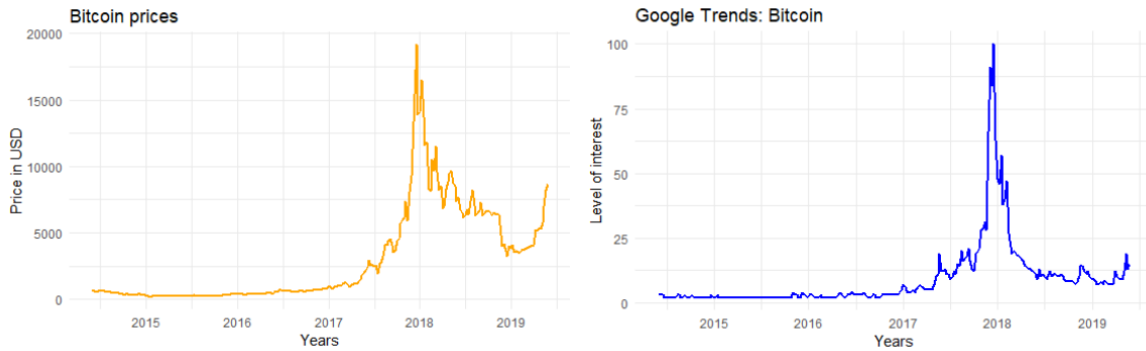


Figura 1 - fonte [5]

Como é possível perceber, o número de pesquisas acompanha o preço do Bitcoin, e pode ser utilizada para prever momentos de euforia e desespero do mercado.

Foi utilizada a média móvel dos últimos 7 dias desse indicador, para que o modelo consiga captar melhor as tendências de movimento, e não sofra grandes influências de movimentos bruscos de um único dia.

Spent Output Profit Ratio (SOPR)

É a Proporção de Lucro nos Saldos Gastos. É calculado a partir dos saldos gastos (*spent outputs*) de uma transação do Bitcoin. Pode ser entendido como o valor realizado dos bitcoins (USD) dividido pelo valor dos bitcoins na criação (USD). Ou simplesmente: Preço de venda / Preço de aquisição. Indica se os *holders* estão vendendo com lucro ou prejuízo, e pode indicar oportunidades de arbitragem, ou se o preço está muito valorizado ou desvalorizado.

Net Unrealized Profit (NUPL)

é representado pela fórmula:

$$NUPL = \frac{(MarketCap - RealizedCap)}{MarketCap}$$

Onde:

- valor de mercado da moeda: é o preço multiplicado pela quantidade de bitcoins disponível
- valor realizado: é calculado avaliando cada unidade de bitcoin individualmente ao preço que foi transacionado pela última vez *on-chain*. pode ser pensado como uma aproximação bruta do custo agregado do bitcoin base, também por vezes referido como o seu "valor armazenado" total.

Indica os resultados líquidos que não foram realizados sobre todo o valor de mercado, podendo indicar sobreavaliação ou subavaliação do ativo.

Coin Days Destroyed (CDD)

"*Coin Days*" são fruto do produto entre certa quantidade de bitcoins e o número de dias em que eles permaneceram na mão de um agente. Coin Days Destroyed avalia não só a quantidade de moedas vendidas (destruídas), mas estabelece peso para quanto maior for o tempo em que elas permaneceram na mão de um agente. Um valor alto para o CDD indica que os *holders* estão abrindo mão de suas posições, logo há uma tendência de instabilidade no mercado.

3. Escolha do Período

Para a escolha do critério utilizado na seleção do período a ser implementado no modelo, buscou-se utilizar o máximo possível de dados que estavam disponíveis. Para tanto, os dados das séries históricas foram obtidos através da plataforma do *Google Trends* e do site *Glassnode*, um site especializado em dados de criptoativos. Assim, por conta de algumas limitações das plataformas e da disponibilidade de dados, o período utilizado foi de **01/02/2011** até **30/07/2021**, utilizando dados **diários**.

Este período foi então dividido em treino e teste. O período de treino foi utilizado para otimizar os parâmetros do modelo sem incorrer em nenhum viés de overfitting ou look-ahead-bias no período de testes. **60%** da amostra de dados foi destinada ao período de treino (01/02/2011 – 17/05/2017 – 1627 dados) e **40%** para o período de teste (18/05/2017 – 30/07/2021 – 1075 dados).

4. Apresentação da Estratégia

A ideia é utilizar um modelo de *machine learning* de classificação binária (alta ou baixa) para prever o comportamento futuro do preço do bitcoin. Foi utilizado um modelo de Árvore de Decisão.

4.1. Algoritmo

O ajuste de uma árvore de decisão ocorre a partir da variável independente que apresenta o maior poder preditivo, subdividindo-se dessa maneira até encontrar uma previsão. Nesse contexto, o poder preditivo significa dividir com base em uma das variáveis independentes, de forma que a variável dependente (alvo) possa ser separada em dois subconjuntos homogêneos bem definidos, obtendo assim o maior ganho de informação possível.

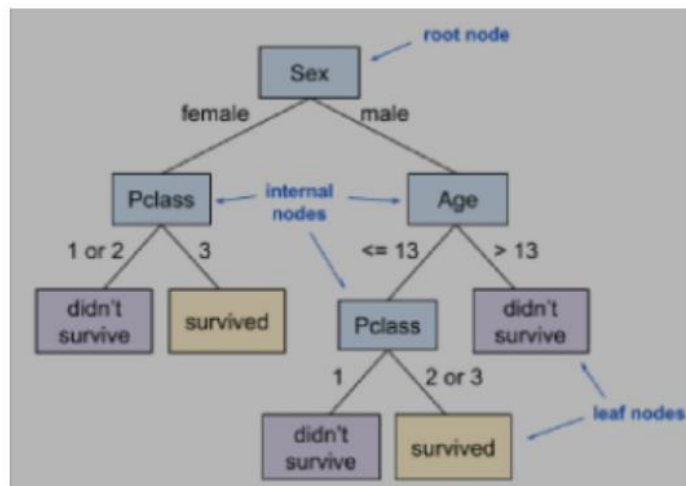


Figura 2: Exemplo de árvore de decisão estimada a partir de um conjunto de treinamento com dados de passageiros do Titanic

Após estimada a árvore de decisão que melhor explica o modelo dadas as condições iniciais (hiperparâmetros) e partindo do pressuposto fundamental de que os dados do conjunto de teste se comportam da mesma maneira que os de treino, podemos prever uma variável dependente a partir dos valores observados para as variáveis independentes naquele dia.

4.2. Hiperparâmetros

São parâmetros ajustáveis definidos a priori que permitem controlar o processo de treinamento de um modelo, e podemos avaliar o desempenho de diferentes modelos contendo combinações distintas de valores de hiperparâmetros para encontrar o melhor. Os hiperparâmetros do modelo utilizado foram otimizados por meio do *GridSearch* (teste exaustivo de diferentes combinações de parâmetros para encontrar aquela que apresenta o melhor resultado). Os parâmetros ótimos obtidos foram:

- Função que define o ganho de informação: impuridade de Gini;
- Profundidade máxima da árvore: 10;
- Número mínimo de amostras presentes em cada folha (decisão): 50;
- Número mínimo de amostras necessárias para fazer uma nova divisão: 5;
- Número máximo de folhas (decisões) geradas pela árvore: 10

5. Considerações sobre o Backtesting

5.1. Look-ahead bias

Para fazer a estimação do modelo sem cair em *look-ahead bias*, o grupo dividiu a estratégia em dois momentos: período de treino (60% dos dados) e período de testes (40% dos dados), onde o ponto de corte definido a priori ocorre no dia 17/05/2017. Além disso, foi utilizado no período de treino o sistema de *cross-validation* para encontrar o conjunto ótimo de hiperparâmetros a serem utilizados pelo modelo.

Vale comentar também que diversos indicadores utilizados, por possuírem um grande histórico, diversos sites já fazem recomendações quanto aos momentos de comprar e vender, por historicamente esses pontos terem se mostrado eficazes (por exemplo, sites indicam que se deve comprar Bitcoin quando o NUPL estiver abaixo de 0% e vender quando estiver acima de 75%). O grupo tomou o cuidado de não inserir nenhuma dessas informações no modelo, para que o algoritmo tivesse que descobrir sozinho os pontos de compra e venda.

5.2. *Overfitting*

O objetivo da otimização de hiperparâmetros é mitigar o *overfitting*, aumentando a capacidade de generalização do modelo com a realização do *pruning* (podagem), onde foram estabelecidas restrições para a profundidade da árvore. Além disso, a divisão da base de dados em treino e teste permite testar os resultados no período de teste sem nenhum tipo de *overfitting*, avaliando como o modelo performaria na vida real.

6. Resultados

Após a realização do processo de treino, e de serem encontrados os parâmetros ótimos para cada indicador, o modelo então foi testado durante o período de 18/05/2017 – 30/07/2021, para avaliar como performaria em um cenário real, sem *overfitting*.

6.1. Retornos

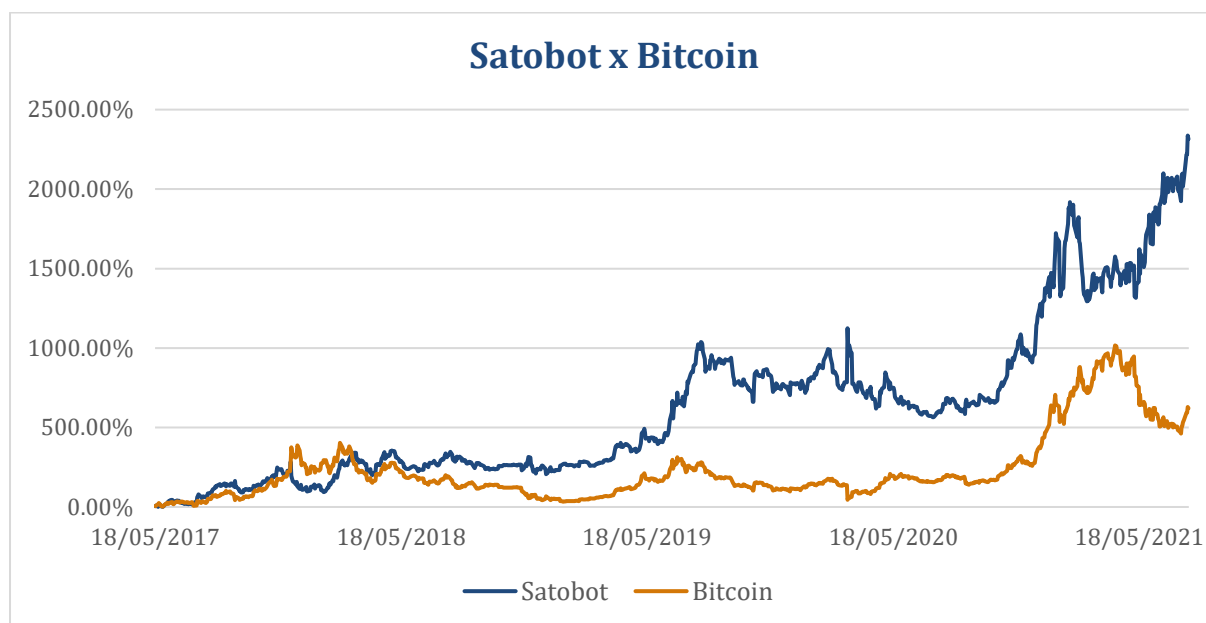


Gráfico 1 – Retornos Acumulados

Como é possível ver no gráfico 1, o Satobot obteve um retorno de **2.314,37%** no período, o que equivale a um retorno de **110,94% a.a.** No mesmo período o Bitcoin obteve um retorno de 621,81%, equivalente a 58,94% a.a. Além disso, é possível perceber que o modelo consegue entregar resultados positivos tanto em momentos de alta como de baixa do Bitcoin.

6.2. Risco

A volatilidade anualizada do modelo foi de **68,90%**, enquanto o Bitcoin apresentou uma volatilidade de 69,03%. Em termos absolutos, é uma volatilidade bem elevada, mas essa é uma característica do mercado de criptomoedas. Além disso, o Satobot conseguiu ficar com uma volatilidade um pouco inferior ao Bitcoin.

Como é possível observar no gráfico 2, o máximo drawdown do modelo foi de -**49,73%**, em janeiro de 2018, contra um máximo drawdown de -73,51% do Bitcoin (nesse mesmo período o drawdown do Satobot foi de apenas -32,81%). A estratégia conseguir apresentar drawdowns inferiores ao Bitcoin em quase todo o período.

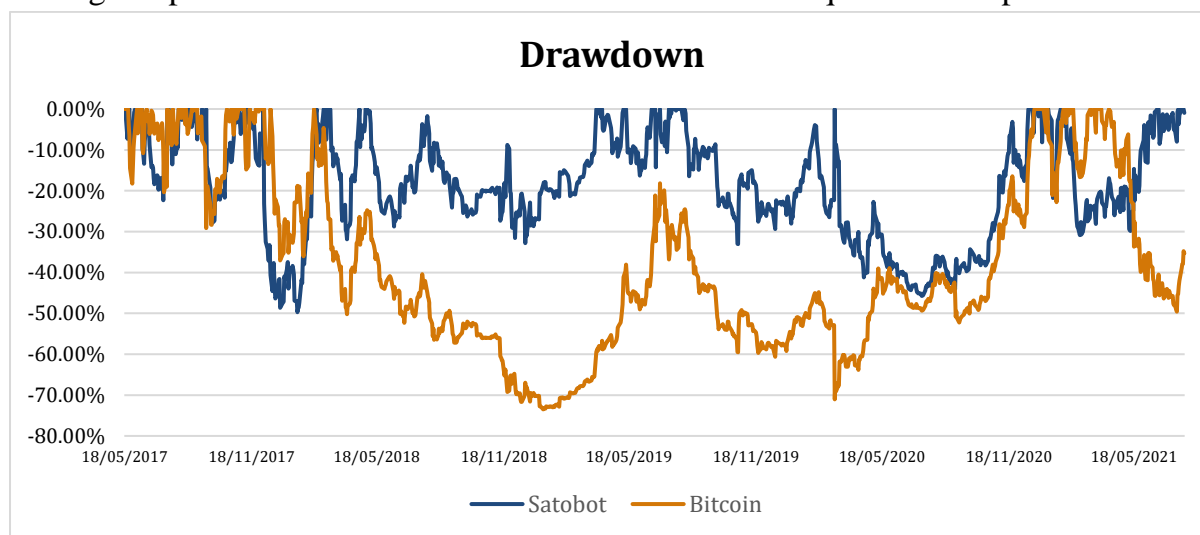


Gráfico 2 – Drawdown observados

6.3. Risco e Retorno

Para avaliar a relação risco e retorno do modelo, utilizamos dois indicadores: Índice de Sharpe e Índice de Calmar. O índice de Sharpe tem como denominador o desvio padrão, enquanto o índice de Calmar tem o drawdown. Para taxa livre de risco, foram utilizados os títulos de 10 anos dos estados unidos (*Tresury Bonds*). Os resultados foram muito bons, indicando boa relação de risco e retorno.

	Satobot	Bitcoin
Índice de Sharpe	1.56	0.81
Índice de Calmar	2.17	0.76

Tabela 1

6.4. Consistência

Quanto ao percentual de acertos, o Satobot apresentou um resultado satisfatório de **52%**, mas ficou abaixo do Bitcoin, que teve 53,30% dos seus retornos positivos. Contudo, o modelo compensa ganhando mais em média quando acerta e perdendo

menos quando erra, como podemos ver na tabela 2. Mesmo acertando menos nas posições vendidas, o modelo conseguiu agregar alfa com as mesmas.

	Satobot	Bitcoin
% Positivos	52.00%	53.30%
% Negativos	48.00%	46.70%
Ganho médio	3.15%	2.97%
Perda Média	-2.60%	-2.79%
% Acerto Compra	53.81%	
% Erro Compra	46.19%	
% Acerto Venda	47.85%	
% Erro Venda	52.15%	

Tabela 2

Além disso, um resultado muito positivo foi a baixa correlação do modelo com o Bitcoin. A correlação do Satobot com o Bitcoin foi de apenas **0,22**, um valor extremamente baixo. Isso indica a capacidade da estratégia de entregar resultados independente do momento de mercado.

7. Conclusão

Ao final dos testes, ficamos muito satisfeitos com os resultados, pois o Satobot conseguiu cumprir muito bem com o seu objetivo de superar o *buy-and-hold* do Bitcoin. Esse resultado ainda é mais significativo quando levamos em consideração que o mesmo conseguiu apresentar retornos superiores, com menor volatilidade e um drawdown muito inferior. Ou seja, o modelo apresentou uma excelente relação risco e retorno, o que é confirmado pelos índices de Sharpe e Calmar. Por fim, a baixa correlação com o ativo demonstra que o modelo tem a capacidade de gerar retornos independente do mercado.

Esses resultados comprovam que é possível obter retornos explorando a ineficiência do mercado, que por diversas vezes se mostra ou sobreavaliado ou com grande desconto, justamente por conta de vieses comportamentais dos investidores. Portanto, estamos confiantes em indicar este modelo na prática.

8. Referências Bibliográficas

- [1] Disponível em: [Service warns you when to buy cryptocurrencies based on Google Trends \(criptoeconomia.com.br\)](https://criptoeconomia.com.br)> acesso em: 04/08/2022;
- [2] Disponível em: [Google, the most popular Bitcoin trend indicator, turns 23 \(cryptonews.net\)](https://cryptonews.net)> acesso em: 04/08/2022;
- [3] Disponível em: [Using Google Trends to Detect Bitcoin Price Bubbles - CoinDesk](https://www.coindesk.com/using-google-trends-to-detect-bitcoin-price-bubbles)> acesso em: 03/09/2022;
- [4] ASLANIDIS, Nektarios; BARIVIERA, Aurelio F.; LÓPEZ, Óscar G. The link between cryptocurrencies and Google Trends attention. *Finance Research Letters*, p. 102654, 2022.
- [5] Disponível em: [AAALbitcoinITISE19.pdf;jsessionid=26B93FA99EAB5152B703D0389E6BD36A \(upc.edu\)](https://www.upc.edu/bitcoinITISE19.pdf;jsessionid=26B93FA99EAB5152B703D0389E6BD36A)> acesso em: 17/08/2022;
- [6] Disponível em: <https://academy.glassnode.com/indicators/sopr/sopr-spent-output-profit-ratio>>
- [7] Disponível em: <https://docs.glassnode.com/basic-api/endpoints/indicators#net-unrealized-profit-loss-nupl>>
- [8] Disponível em: <https://hackernoon.com/how-to-use-google-trends-api-with-python>>
- [9] Disponível em: <https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html>>
- [10] CERQUEIRA, V.; TORGO, L.; MOZETIC, I. Evaluating time series forecasting models - an empirical study on performance estimation methods. 2019
- [11] Disponível em: <https://insights.glassnode.com/the-week-onchain-week-30-2022/>>
- [12] Disponível em: <https://coinmetrics.io/on-chain-indicators/>
- [13] Han, Weihao, David Newton, Emmanouil Platanakis, Charles Sutcliffe, and Xiaoxia Ye. "Cryptocurrency factor portfolios: Performance, decomposition and pricing models." *Decomposition and Pricing Models* (2021).
- [14] Renkens, M. "Quantitative Factor Investing Strategies in the Cryptocurrency market." (2018).
- [15] Micaletti, Raymond. "Relative Sentiment and Machine Learning for Tactical Asset Allocation." *Available at SSRN 3475258* (2019).
- [16] Spilak, Bruno. "Deep neural networks for cryptocurrencies price prediction." Master's thesis, Humboldt-Universität zu Berlin, 2018.