

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ECONOMIA**

Pedro Luiz Caixeta Rodrigues

Investigações sobre o impacto do halving no preço do bitcoin

Juiz de Fora
2023

Pedro Luiz Caixeta Rodrigues

Investigações sobre o impacto do halving no preço do bitcoin

Monografia apresentada pelo acadêmico Pedro Luiz Caixeta Rodrigues ao curso de Ciências Econômicas da Universidade Federal de Juiz de Fora, como requisito parcial para a obtenção do título de Bacharel em Ciências Econômicas.

Orientador: Paulo César Coimbra Lisbôa

Juiz de Fora
2023

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Caixeta Rodrigues , Pedro Luiz .
Investigações sobre o impacto do halving no preço do bitcoin /
Pedro Luiz Caixeta Rodrigues . -- 2023.
55 p. : il.

Orientador: Paulo César Coimbra Lisbôa
Trabalho de Conclusão de Curso (graduação) - Universidade
Federal de Juiz de Fora, Faculdade de Economia, 2023.

1. Ativos digitais . 2. Finanças . 3. Modelos de previsão e
classificação . 4. Bitcoin . 5. Halving . I. Coimbra Lisbôa, Paulo César
, orient. II. Título.



UNIVERSIDADE FEDERAL DE JUIZ DE FORA
REITORIA - FACECON - Depto. de Economia

FACULDADE DE ECONOMIA / UFJF

ATA DE APROVAÇÃO DE MONOGRAFIA II (MONO B)

NA DATA DE 13/07/2023, A BANCA EXAMINADORA, COMPOSTA PELOS PROFESSORES

1 – PAULO CÉSAR COIMBRA LISBÔA - ORIENTADOR; E
2 – ROGERIO SILVA DE MATTOS,

REUNIU-SE PARA AVALIAR A MONOGRAFIA DO ACADÊMICO PEDRO LUIZ CAIXETA
RODRIGUES,
INTITULADA: INVESTIGAÇÕES SOBRE O IMPACTO DO *HALVING* NO PREÇO DO
BITCOIN.

APÓS PRIMEIRA AVALIAÇÃO, RESOLVEU A BANCA SUGERIR ALTERAÇÕES AO
TEXTO APRESENTADO, CONFORME RELATÓRIO SINTETIZADO PELO ORIENTADOR.
A BANCA, DELEGANDO AO ORIENTADOR A OBSERVÂNCIA DAS ALTERAÇÕES
PROPOSTAS, RESOLVEU APROVAR A REFERIDA MONOGRAFIA.



Documento assinado eletronicamente por **Paulo César Coimbra Lisbôa, Professor(a)**, em 15/07/2023, às 16:01, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Rogério Silva de Mattos, Professor(a)**, em 15/07/2023, às 17:10, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1367656** e o código CRC **ADA2BBE7**.

RESUMO

O *halving* é um evento que ocorre a cada 4 anos, no qual é estabelecida uma certa quantidade de bitcoins a serem criadas, sendo que essa quantidade é reduzida pela metade em relação ao período anterior. Tendo em vista esse evento, esta monografia teve como objetivo investigar o impacto do halving no preço do *Bitcoin*. A priori foi efetuada uma revisão da literatura referente a essa criptomoeda e a apresentação de conceitos importantes inerentes ao tema como a *BlockChain* e a mineiração, para possibilitar um conhecimento dos mecanismos de funcionamento e diferenças em relação às moedas fiduciárias. Ademais, para obter evidências robustas do objeto de pesquisa realizado, foram utilizados modelos de classificação e regressão, sendo esses, regressão logística, Gradient Boosting e Random Forest, sendo o modelo com melhor desempenho o de regressão logística. Também foi utilizada a metodologia SARIMAX com a finalidade de fazer previsão com um modelo de projeção univariada. Constaram-se fortes evidências de um impacto direto do halving no preço do Bitcoin, e possibilidades de alta dos preços em um horizonte próximo.

ABSTRACT

The halving is an event that occurs every 4 years, in which a certain amount of bitcoins to be created is established, and this amount is reduced by half compared to the previous period. Taking into account this event, the aim of this monograph was to investigate the impact of halving on the price of Bitcoin. Initially, a literature review was conducted on this cryptocurrency and the presentation of important concepts related to the topic, such as Blockchain and mining, to enable an understanding of the mechanisms of operation and differences compared to fiat currencies. Furthermore, in order to obtain robust evidence for the research objective, classification and regression models were used, including logistic regression, Gradient Boosting, and Random Forest, with logistic regression performing the best. The SARIMAX methodology was also used to forecast using a univariate projection model. Strong evidence was found of a direct impact of halving on the price of Bitcoin and possibilities of price increases in the near future.

LISTA DE ILUSTRAÇÕES

Figura 1 - Sistema de verificação da autenticidade e propriedade.....	12
Figura 2 - Consumo de energia utilizado ao longo dos anos na mineração de Bitcoins	18
Figura 3 - Demonstração de como o halving interfere nos ganhos dos mineradores	20
Figura 4 - Expansão da base monetária do Bitcoin.....	20
Figura 5 - Gráfico em escala logarítmica da linha do tempo das cotações do Bitcoin em relação com o evento do Halving	26
Figura 6 - Gráfico de Candles das cotações do Bitcoin em relação com o evento do Halving	27
Figura 7 - Recortes dos Halvings anteriores dos anos de 2016 e 2020	28
Figura 8 - Relatório de classificação dos modelos	30
Figura 9 - Matrizes de confusão dos diferentes modelos utilizados	31
Figura 10 - Curvas ROC dos modelos elaborados.....	32
Figura 11 – Primeira diferenciação da série temporal	33
Figura 12 – Segunda diferenciação da série temporal	33
Figura 13 - FAC e FACP	35
Figura 14 – Modelo Estimado.....	35
Figura 15 - – Decomposição STL.....	37
Figura 16 - – Previsão do modelo SARIMAX	37

LISTA DE ABREVIATURAS E SIGLAS

BTC	Bitcoin
P2P	Peer-to-Peer

SÚMARIO

1 INTRODUÇÃO	5
2 OBJETIVOS DA PESQUISA	7
2.1 OBJETIVOS GERAIS	7
3 REVISÃO DA LITERATURA	9
3.1 A ORIGEM DA MOEDA E SEU DESENVOLVIMENTO	9
3.2 O QUE É O BITCOIN?	11
3.3 O QUE É A BLOCKCHAIN?	13
3.4 O BITCOIN E SUA COMPARAÇÃO COM AS MOEDAS FIDUCIÁRIAS	14
3.5 O QUE É A MINERAÇÃO?	17
3.6 O QUE É O HALVING?	19
4 METODOLOGIA	21
4.1 Base de Dados	25
4.2 EVIDÊNCIAS EMPÍRICAS DA INFLUÊNCIA DO HALVING NAS VARIAÇÕES DOS PREÇOS DO BITCOIN	25
5 – RESULTADOS	29
6 – Considerações finais	39
REFERÊNCIAS	41
APÊNDICE A – Código em Python das regressões	43
APÊNDICE B – Modelo de previsão SARIMAX	49

1 INTRODUÇÃO

O *Bitcoin* ainda é um ponto de inflexão para a maior parte dos economistas e dos indivíduos envolvidos em mercados financeiros. Nassim Taleb, segundo matéria do Valor Econômico (2022) apresenta uma diversidade de críticas para esse ativo, o classificando como uma doença contagiosa, seja por conta da extrema volatilidade que esse criptoativo tem como característica, a falta de um valor intrínseco atribuído a moeda digital, assim como a ausência de um órgão que o regule, a possibilidade da utilização desse ativo para fraudes fiscais e sonegações, e por fim, pode-se destacar os danos ao meio ambiente causados devido ao grande gasto de energia elétrica utilizado para a mineração desses ativos, visto que, os mineradores, em sua maior parte buscam países que ofereçam energia elétrica mais barata, que possuem como fonte energética, a queima de combustíveis fósseis e carvão, essas que são extremamente poluentes e danosas ao meio ambiente.

Contudo, algo que não pode ser descartado é que de fato, o bitcoin surge com uma potencial forma de pagamentos disruptiva, seja por conta da tecnologia da *blockchain*, seja pelo fato de também ser visto como uma potencial forma revolucionária de criar uma sociedade mais justa baseada nas plataformas abertas e dados abertos (BAUR, et al., 2015). A possibilidade de realizar transações online com baixíssimos custos de transação, praticamente descartam eventuais dificuldades que possam aparecer em relação a câmbio e território, trazendo a perspectiva de bitcoin como o “futuro da moeda” (KAMINSKA, et al., 2016).

No final do ano de 2020 e início do ano de 2021, as altas crescentes do Bitcoin em níveis colossais chamaram atenção de diversos especialistas e estudiosos da área, que buscavam entender os motivos dessa alta, e apesar de muitos o classificarem como uma bolha especulativa, é preciso destacar que no dia de 11 de Maio de 2020, ocorreu o *halving*, isto é, um evento que acontece a cada 4 anos, na qual o número de Bitcoins pago aos “mineradores” (indivíduos responsáveis por criptografar as transações realizadas com bitcoin ao livro razão público desse criptoativo, na qual é comumente referida como “Blockchain”) é reduzido, e historicamente após o *halving*, as cotações do bitcoin tendem a aumentar. Tomando como janela as cotações do bitcoin de 1 ano antes a 1 ano após o *halving*, é possível ver que os preços dispararam, contudo, essa alta vem registrando taxas cada vez maiores.

Por conta disso, esta monografia tem por objetivo geral investigar se é possível estabelecer uma relação existente entre o halving e o preço do bitcoin. Também deverá ser introduzido nessa monografia, uma explicação sobre o que é o Bitcoin, evidenciando suas principais características, assim como trazer um pouco mais de fundamentação ao halving evidenciando por qual motivo ele ocorre, e sua importância para que o sistema *peer-to-peer*, isto é, uma arquitetura de redes de computadores em que cada um dos pontos ou nós da rede funciona tanto como cliente quanto como servidor, permitindo o compartilhamento de serviços e dados sem a necessidade de um servidor central encontre-se em pleno funcionamento.

Além da presente introdução, as seções estão organizadas da seguinte forma: seção 2, serão apresentados os objetivos gerais e específicos da tese. Na seção 3 será apresentado o arcabouço e referencial teórico utilizados para a construção da tese, aprofundado os conceitos de moeda, o surgimento do bitcoin, destacando suas principais características e sua diferenciação frente as moedas fiduciárias, assim como as principais variáveis que influenciam a sua cotação. Dessa forma são elaboradas as seguintes subseções, 3.1 - A origem da moeda e seu desenvolvimento; 3.2 - O que é o bitcoin, 3.3 - O que é a Blockchain; 3.4 - O Bitcoin em comparação com as moeda fiduciárias; 3.5 - O que é a mineração e 3.6 – O que é o Halving. Na seção 4 será trabalhado a metodologia utilizado na pesquisa, onde são apresentadas as evidências empíricas, a base de um modelo matemático heurístico e o modelo de regressão logística, também está presente o desenvolvimento da seção 4.1 na qual são apresentadas evidências empíricas da influência do halving no comportamento de alta do bitcoin

2 OBJETIVOS DA PESQUISA

Esta Monografia tem como objetivo investigar o impacto do halving no preço do bitcoin, visto que, há muitos questionamentos em comunidades acadêmicas e mercados financeiros, sobre quais seriam os determinantes da sua cotação, além disso, existem teses de que a especulação seria a principal variável responsável por alavancar seus preços. Por conta disso, dada a presença de uma forte correlação entre o evento do halving, na qual a emissão de novos bitcoins é cortado pela metade e há altas históricas desse ativo digital, desperta-se um ímpeto em analisar se de fato essas variáveis estão correlacionadas. Outrossim, constatada ou não a correlação e causalidade entre essas duas variáveis, já nos permite trazer maior previsibilidade e confiabilidade para investidores e agentes dos mercados financeiro, pois o próximo halving está previsto para acontecer no ano de 2024 e desperta muita expectativa sobre se essa tendência de alta será mantida.

2.1 OBJETIVOS GERAIS

Desse modo, após apresentados os principais conceitos que envolvem a temática, cabe salientar que esta pesquisa tem como objetivo mensurar e determinar o papel do halving como determinante no que tange as variações nos preços do Bitcoin. A principal base de dados utilizada será a cotação do bitcoin ao longo dos anos, e as datas nas quais ocorreu o evento do halving. Dessa forma, haverá a possibilidade de afirmar se de fato o halving é uma variável não exógena ao comportamento de alta do Bitcoin, visto que, empiricamente é possível constatar que durante os períodos de ocorrência deste evento, as cotações deste ativo digital dispararam, o que indica um forte viés de correlação entre essas duas variáveis.

2.2 OBJETIVOS ESPECÍFICOS

Será realizada uma comparação entre os modelos de regressão logística Random Forest e Gradient Boosting, dessa forma, irá buscar compreender se esse viés de correlação e causalidade entre o evento do Halving e a cotação do Bitcoin não se trata de uma correlação espúria, isto é, duas variáveis que inicialmente apresentam algumas evidências de viés de correlação, porém são independentes entre si, e esse

viés só ocorria por coincidência, não implicando em causalidade, isto é sua variável explicativa sendo significativa no que tange sua estimação da sua variável dependente. Também será utilizada de uma estimação univariada, com o objetivo de verificar o quanto um modelo autoprojeto consegue prever a série de forma satisfatória, tendo em vista que modelos baseados na metodologia de Box-Jenkins são comumente usados para preços de ações. Nesse viés, é interessante analisarmos se em relação a uma série de preços do Bitcoin, que possui diversas características de outros ativos financeiros, essa metodologia consegue ser efetiva.

De forma concomitante, caso seja constatada a relação entre as duas variáveis, também terá o objetivo de fornecer maior previsibilidade e segurança aos entusiastas e investidores do criptoativo sobre seu comportamento de alta e baixa ao longo do tempo.

3 REVISÃO DA LITERATURA

Este capítulo faz uma revisão dos conceitos apresentados no que tange aos temas origem e evolução da moeda, além das características necessárias para determinado objeto de estudo, poder ser classificado como moeda.

Outro objetivo desta seção é apresentar os conceitos do Bitcoin e Blockchain explicando sua origem e principais características, assim como o halving, na qual foi elucidado por qual motivo ele ocorre e sua importância para o funcionamento do sistema peer-to-peer.

3.1 A ORIGEM DA MOEDA E SEU DESENVOLVIMENTO

Inicialmente, as sociedades primitivas não se utilizavam de moedas para realizar trocas de mercadorias entre os indivíduos. Sem a moeda, há uma economia que funcionaria na base do escambo, no entanto, provou ser uma maneira inapropriada de distribuir esses recursos, uma vez que exigia a dupla coincidência de anseios (MANKIW,2014). Além disso, as transações que esse sistema permitia eram mais simples e não conseguiam abranger a quantidade de transações indiretas e mais complexas presentes em sociedades mais desenvolvidas. Assim como destacado por Menger:

[...]Mesmo no caso relativamente simples e muitas vezes recorrente, onde uma unidade econômica, A, requer sobre a uma commodity pertencente à B, e B requer uma possuída por C, enquanto C quer uma que é propriedade de A — mesmo aqui, sob uma regra de mera permuta, a troca dos bens em questão seria, em regra, inevitavelmente deixada por fazer. (MENGER 2011, p.15-16):

Por conta dessa necessidade em encontrar algo de valor comum para os indivíduos e que pudesse ser utilizado como um meio de troca, surgem as moedas-mercadoria, que possuíam algum valor intrínseco, pois poderiam ser utilizadas para diversos propósitos, entre um deles, ser usado como dinheiro. Foram utilizados diversos tipos de moedas-mercadoria ao longo da história, contudo, quanto mais amplas se tornavam as redes comerciais, se tornava mais necessário a confiança das partes envolvidas no valor dessa moeda, é que essa se caracterizasse uma unidade

de valor, que facilita a avaliação e o cálculo; em um recipiente de valor, que permite que as transações econômicas sejam conduzidas durante longos períodos e também a despeito das distâncias geográficas (FERGUSON, 2009).

Nesse mesmo viés, outras características que passaram a ser desejáveis para essas moedas-mercadorias é de que fossem duráveis, fungíveis, portáteis e confiáveis. Como preenchem a maioria desses critérios, ao longo dos milênios os metais, como ouro, prata e bronze foram considerados como a matéria-prima monetária ideal (FERGUSON, 2009).

Com o passar dos séculos, ocorre uma evolução das moedas-mercadorias para as moedas fiduciárias, inicialmente pela dificuldade de se medir a pureza do ouro, ou a dificuldade que pode ser carregar barras de ouro para determinadas negociações, passa-se a utilizar moedas cunhadas em ouro, em seguida podem ser utilizados notas de papel que cumprem o papel de resgate de determinada quantidade de ouro nos bancos, com o tempo o próprio lastro em ouro se torna insignificante (MANKIWI, 2014), visto que as pessoas não se importam mais em resgatar essa quantidade em ouro dos bancos, porém a confiança das pessoas que este papel moeda possua algum valor é um fator inexorável para o funcionamento do sistema.

De forma complementar, com o passar dos anos se desenvolveu que uma instituição governamental fosse responsável por determinar a quantidade de moeda fiduciária circulante na economia, essa instituição é o Banco Central. Para Samuelson e Nordhaus (2012), a política monetária é uma das ferramentas mais importantes para estabilizar os ciclos econômicos. De acordo com Corazza (2007) os bancos centrais, em sua concepção moderna, não foram criados arbitrariamente pelo Estado, mas representam o produto necessário da própria evolução do sistema bancário privado submetido a crises recorrentes.

Vale destacar que ao longo da história, diversos bancos assumiram o papel de um Banco Central, mesmo que essa concepção ainda não fosse tão bem definida. O Riksbank da Suécia, fundado em 1668, embora fosse uma instituição privada, tinha seus gestores escolhidos diretamente pelo rei, além disso, era autorizado a emitir novas notas, informação disponível na página oficial do banco Riksbank (2019). Já o Monte dei Paschi di Siena, fundado em 1472, tinha atuação semelhante a um fundo de caridade, além de oferecer o serviço de custódia da fortuna aos nobres, informação que também pode ser encontrada na página oficial do banco Monte dei Paschi di Siena (2017).

Entretanto, para Hayek (2017) não foi necessário um planejamento central para que a moeda fosse introduzida na sociedade, sendo essa introduzida de forma espontânea na sociedade como fruto da percepção dos indivíduos da utilidade e necessidade de um meio universal de troca. Para o autor, moedas de curso legal são indesejáveis para a sociedade como um todo, pois ao dar ao governo o poder de regular e emitir dinheiro, gerava uma série de incertezas e arbitrariedades, propondo então uma sociedade na qual empresas privadas pudessem emitir moedas livremente, pois dessa forma, a concorrência entre essas empresas seria responsável por oferecer produtos monetários com maior qualidade do que as moedas fiduciárias.

De toda forma, hodiernamente os bancos centrais tem um papel decisivo para a sociedade na qual está relacionada, visto que são responsáveis por definir a base monetária de uma economia e a oferta monetária em uma economia, influenciando diretamente o nível geral de preços e o poder de compra dos indivíduos dessas economias.

3.2 O QUE É O BITCOIN?

Criado em 2008 pelo programador anônimo Satoshi Nakamoto, com o objetivo de “desenvolver uma versão puramente peer-to-peer de dinheiro eletrônico, permitindo que pagamentos on-line fossem enviados diretamente de uma parte para outra, sem passar por uma instituição financeira ou um agente intermediário” (SATOSHI, 2008, p.1).

Não obstante, sem a presença de um agente intermediário, surge uma necessidade de criar um mecanismo que pudesse evitar o fenômeno do “gasto duplo”, isto é, a utilização das mesmas moedas digitais mais de uma vez. De acordo Satoshi Nakamoto

O problema, claro, é o sacador não poder confirmar se um dos pagadores não fez gasto duplo da moeda. Uma solução comum é a introdução de uma autoridade central confiável, ou casa da moeda, que verifique o gasto duplo para todas as transações. Depois de cada transação, a moeda deve ser devolvida à casa da moeda para a emissão de uma nova, e apenas moedas emitidas diretamente da casa da moeda são confiáveis de não ser gastas duplamente. O problema desta solução é que o destino de todo o sistema monetário depende da empresa que gerencia a casa da moeda, com todas as transações tendo de passar por ela, assim como um banco. (SATOSHI, 2008, p.2)

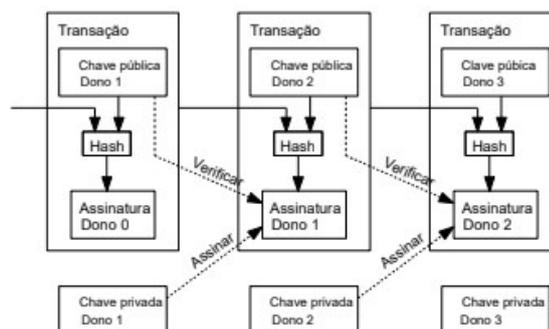
Por conta disso, foi desenvolvido o sistema da *Blockchain*, isto é, uma espécie de livro diário na qual estão listados os históricos de todas as transações realizadas, Novas transações são verificadas contra o blockchain de modo a assegurar que os mesmos bitcoins não tenham sido previamente gastos, eliminando assim o problema do gasto duplo (ULRICH, 2014).

Ainda de acordo com Nakamoto:

[..]Para resolver isso, propusemos uma rede peer-to-peer usando prova de trabalho para gravar um histórico público de transações que rapidamente se torna computacionalmente impraticável [...] A rede insere data e hora nas transações através de um hash em uma cadeia contínua de prova-de-trabalho à base de hash, formando um registro que não pode ser alterado sem refazer a prova-de-trabalho. (SATOSHI, 2008, p.8)

Ou seja, ao se realizar a compra/venda de um bitcoin, são gerados chaves públicas e privadas para os indivíduos envolvidos nessa transação, as chaves privadas possuem como objetivo ser uma espécie de senha, o detentor dessa chave é o único capaz de movimentar aqueles ativos, nesse viés uma expressão que se tornou bastante conhecida entre os entusiastas de criptomoedas que ajuda a sintetizar essa ideia é “not your Keys, not your coins” em tradução livre, seria algo semelhante a “se não possui as chaves, não são as suas moedas”, visto que, essas chaves privadas são irrecuperáveis e intransferíveis, ou seja, em caso de perda destas por parte do proprietário, essas moedas estariam perdidas e incapazes de realizar novas transações. Já as chaves públicas, tem o objetivo de serem um registro público da transferência de propriedade daquelas bitcoins. Em geral, qualquer indivíduo pode verificar as chaves públicas de outrem, para constatar se de fato, tal transação foi realizada e quais as partes envolvidas, contudo apenas o detentor da chave privada, podem movimentar aquela quantia em bitcoins.

Figura 1 - Sistema de verificação da autenticidade e propriedade



Fonte: Bitcoin: A Peer-to-Peer Electronic Cash System (2008)

Dessa forma, também é possível entender melhor o papel dos “mineradores” para o funcionamento dos pagamentos realizados com Bitcoin, sem a presença de agentes intermediários, esses agentes, “emprestam” o hardware de seus computadores, para que a rede Bitcoin possa verificar a autenticação das transações realizadas, o primeiro minerador que consegue registrar uma transação na Blockchain, é recompensado com um pagamento de Bitcoins. A criptografia de chave pública garante que todos os computadores na rede tenham um registro constantemente atualizado, o que impede o gasto duplo e qualquer tipo de fraude. (ULRICH, 2014)

Ao resolver o problema do gasto duplo, dificuldades enfrentadas por outras criptomoedas anteriores como o E-Gold, Webmoneu e Liberty Reserve (COINDESK, 2015), o Bitcoin torna-se, não apenas uma moeda digital, mas sim o primeiro sistema de pagamentos on-line em escala global de forma totalmente descentralizada e sem a presença de nenhum intermediador.

3.3 O QUE É A BLOCKCHAIN?

Tecnologia desenvolvida em 2008 por Satoshi Nakamoto para que pudesse viabilizar a utilização do Bitcoin como meio um meio de pagamento seguro e a prova de fraudes, a blockchain atua como um livro-razão público, registrando todas as transações realizadas em determinado mercado, permitindo assim que um qualquer indivíduo possa verificar a autenticidade dessas transações.

O funcionamento da Blockchain se dá por forma de um protocolo de registro distribuído onde as informações são armazenadas em blocos conectados. Essencialmente são transformados os dados da uma transação realizada em *hash*, isto é, um código de números e letras que simplifica as informações dos dados iniciais, caso seja feita uma nova transação, o novo hash formará um novo bloco que também contém o hash anterior, e assim sucessivamente, cada nova transação realizada torna a rede cada vez mais segura.

Nakamoto (2008) descreve da seguinte forma o funcionamento da rede:

- 1) Novas transações são transmitidas para todos os nós.
- 2) Cada nó coleta novas transações em um bloco.
- 3) Cada nó trabalha para encontrar uma difícil prova-de-trabalho para o seu bloco.
- 4) Quando um nó encontra uma prova-de-trabalho, ele transmite o bloco para todos os nós.

- 5) Os nós aceitam o bloco somente se todas as suas transações são válidas e já não foram gastas.
- 6) Os nós expressam sua aceitação do bloco, trabalhando na criação do próximo bloco na cadeia, usando o hash do bloco aceito como o hash anterior. (SATOSHI, 2008, p.3)

Por fim, a tecnologia da Blockchain, apesar de ter sido inicialmente desenvolvida com a única finalidade de possibilitar o funcionamento do Bitcoin, atualmente cumpre uma função totalmente disruptiva, podendo ser creditada como uma das criações mais importantes do século, visto que, pode ser utilizada para uma infinidade de processos além da negociação de criptomoedas, como a validação de contratos, transações financeiras tradicionais, vendas de criptoartes, como as NFTs por exemplo, entre outra infinidade de serviços, garantindo maior privacidade, segurança e eficiência as partes envolvidas

3.4 O BITCOIN E SUA COMPARAÇÃO COM AS MOEDAS FIDUCIÁRIAS

Ao perguntar para um economista, quais as finalidades principais que uma moeda deve apresentar, a resposta mais provável seria acerca de um bem de valor comum, que pudesse ser utilizado como reserva de valor, isto é, um meio de transferir o poder de compra do presente para o futuro (MANKIWI, 2014). E que também possa ser utilizada como unidade de conta, ou seja, a moeda define os termos segundo os quais os preços são determinados e as dívidas registradas. A moeda constitui o padrão por meio do qual são mensuradas as transações econômicas (MANKIWI, 2014). Isso permite que bens diferentes sejam comparados entre si. Por fim, essa que talvez seja a função mais essencial da moeda, é a sua utilização como meio de troca, isto é, algo de comum valor entre duas partes, e que pode ser utilizado para adquirir bens e serviços. Como evidenciado anteriormente, sem uma moeda que possa ser utilizada como meio de troca, o comércio passa a exigir a dupla coincidência de anseios, ficando à mercê da improvável eventualidade de dois indivíduos, cada um deles com um determinado bem que o outro deseja, se encontrarem no momento exato, no lugar exato, para realizarem esse intercâmbio (MANKIWI, 2014).

Portanto, para David Yermack (2013), dadas as características e finalidades de uma moeda, não podemos classificar o bitcoin como tal, e sim, como um investimento especulativo, visto que enfrenta dificuldades para cumprir essas funções,

principalmente no que tange o seu papel como reserva de valor e unidade de conta, pois além da alta volatilidade, que o inibe de ser uma reserva de valor confiável, os preços dos bens em Bitcoins são caracterizados pelas várias casas decimais a esquerda, o que dificulta sua utilização no dia-a-dia.

Ainda de acordo com Yermack:

Para que o bitcoin se torne mais do que uma curiosidade e se estabeleça como uma moeda de boa-fé, seu valor diário precisará se tornar mais estável para que possa servir de forma confiável como reserva de valor [...] Conforme descrito acima, o bitcoin também enfrenta dificuldades devido ao seu preço decimal não ortodoxo de mercadorias, a escassez de comerciantes que as aceitam e o complicado processo de aquisição bitcoins de um fornecedor, entre outras questões. (YERMACK, 2013)

Contudo, no que tange sua utilização como meio de troca, assim como destacado por (BAUR, et al., 2015), por mais que ainda apresente algumas barreiras, como por exemplo, dificuldade de usabilidade, riscos de volatilidade e segurança, em geral, a sua utilização vem ganhando bastante força dado as baixas taxas de transação e o potencial alcance de aceitação internacional. Gigantes do ramo da tecnologia, como a Microsoft, já aceitam pagamentos em Bitcoins por certos serviços de assinatura e produtos.

Um fato curioso que vale se destacar, é a importância da data de 22 de dezembro de 2010 para os entusiastas do bitcoin, pois trata-se do momento em que foi registrado a primeira transação envolvendo a moeda, isto é, sua utilização como um meio de troca, que até aquele momento, não possuía sequer um equivalente em preço. O então programador Laszlo Hanyecz, um dos primeiros mineradores de criptomoeda que se tem registro até então, comprou duas pizzas pelo valor de 10.000 Bitcoins, algo em torno de 367 milhões de dólares, na cotação atual (cotação na data de 21/01/2022), assim como destacado pela revista Exame (2021)

Outra característica que distancia bastante o Bitcoin das moedas tradicionais é seu comportamento deflacionário, visto que, existe uma oferta limitada de moedas, algo que não ocorre com as moedas fiduciárias, que podem ser emitidas novas unidades a qualquer momento, contudo, gerando efeitos no índice de preços. A quantidade máxima de bitcoins que podem ser mineiradas é de 21 milhões de

unidades, portanto os Bitcoins que se encontram impossibilitados de serem transacionados, dado fatores como usuários que perderam o acesso as suas chaves privadas, estão permanentemente inutilizados. A consequência, por ficarem “fora” de circulação, seria um aumento no poder de compra do restante de bitcoins existentes (ULRICH, 2014). Isso significa, que se a demanda pelo Bitcoin aumenta e caso a grande parte das pessoas o aceite como um meio de troca, o nível geral de preços diminuiria, e a quantidade de bens e serviços que um indivíduo conseguiria comprar com a mesma quantidade Bitcoins tenderia a aumentar, algo contrário ao que ocorre com as moedas mais tradicionais, como o Real, Euro e Dólar. Uma vez que um número predeterminado de moedas já entrou em circulação, o incentivo pode migrar inteiramente para taxas de transação e ser completamente livre de inflação (SATOSHI,2008)

Para Reisman (2010, p.1), “uma moeda que se aprecia ao longo do tempo com certeza não representa nenhuma ameaça à saúde de uma economia”. Contudo, há muitas discordâncias entre os economistas sobre o quão desejável é o efeito deflacionário em uma economia, pois em um cenário de inflação zero ou abaixo disso, o poder de atuação do Banco Central é reduzido, deixando essas economias mais voláteis. Blanchard (2011), destaca que a deflação limita a capacidade de uma política monetária de afetar o produto. além disso, pode gerar efeitos negativos no crescimento e na renda, visto que, se os agentes tem expectativas que os preços irão cair ainda mais no futuro, provavelmente irão postergar o consumo, por sua vez, as empresas com margens apertadas irão reduzir sua produção, visto que, haverá uma relutância em reduzir os salários nominais, assim como destacado por Mankiw (2014):

As empresas relutam em cortar os salários nominais de seus trabalhadores, e os trabalhadores relutam em aceitar tais cortes. Um corte de 2% no salário, em um mundo com inflação zero, é, em termos reais, a mesma coisa que um aumento de 3% com uma inflação de 5%; mas os trabalhadores nem sempre têm essa percepção. (MANKIW, 2014, p.144)

Portanto, apesar do Bitcoin conseguir solucionar alguns efeitos negativos da inflação, como por exemplo, o efeito Cantillon, isto é, quando alterações na quantidade de moeda podem causar vários tipos diferentes de efeitos reais sobre o setor real da economia, dependendo de quem primeiro recebeu o dinheiro (IORIO, 2014). Ou seja, aqueles que primeiro recebem a moeda recém-criada, não sofrerão de forma tão intensa, os efeitos da desvalorização de poder de compra quanto aqueles que

receberão por último essa moeda, intensificando as desigualdades causadas pelo imposto inflacionário, o bitcoin também pode agir como um vetor para diminuir crescimento, caso fosse o principal meio de troca de determinada economia, dificultando sua aceitação como moeda corrente para a maioria dos países.

Por fim, vale destacar que diferentemente das moedas tradicionais, o Bitcoin não possui nenhum agente controlador ou regulamentador, dado o fato que uma das razões que ajudaram a impulsionar e popularizar essa criptomoeda, foi a falta de confiança crescente no sistema financeiro de reservas fracionárias, e na condução da política monetária por parte dos governos centrais após a crise de 2008 e a falência do banco Lehman Brothers. (ULRICH, 2014)

3.5 O QUE É A MINERAÇÃO?

A mineração de Bitcoins é o único modo possível de colocar novas moedas em circulação. Ao realizar o registro e validação das transações realizadas com Bitcoin na Blockchain, os mineradores utilizam de seus computadores para calcularem milhares funções hash, tanto para criar novas moedas, como para verificar a autenticidade de transações já realizadas, apenas o primeiro “minerador” a decodificar uma dessas funções recebe o prêmio de um pagamento em Bitcoins. Por conta disso, o processo de mineração torna-se caro, visto que necessita de computadores de alta potência e também há um grande gasto de energia. Além disso, trata-se de uma atividade pouco recompensadora, pois apenas um player é o vencedor entre os diversos envolvidos na tentativa de validar os blocos. Vale ressaltar que há evidências significativas de uma relação entre a atividade mineiradora e o preço do criptoativo (CAMACHO,2019)

O pagamento do prêmio aos mineradores cumpre um papel essencial para garantir a integridade do sistema, assim como descrito por Satoshi Nakamoto (2008):

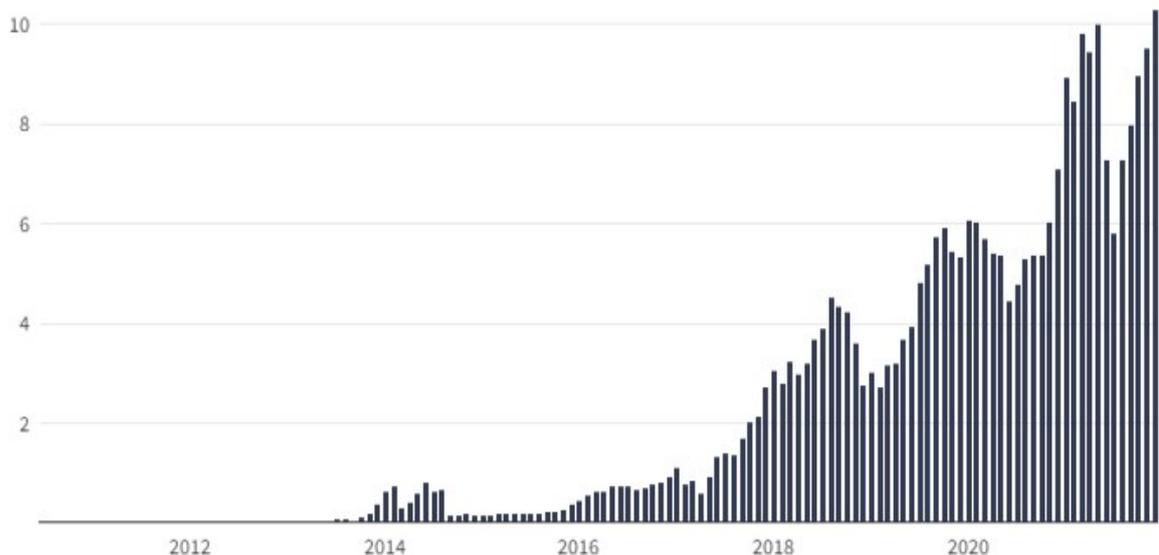
O incentivo deve encorajar os nós a se manterem honestos. Se um player ganancioso é capaz de reunir mais poder de CPU do que todos os honestos, ele teria que escolher entre usá-lo para defraudar as pessoas roubando seus pagamentos, ou usá-lo para gerar novas moedas. Ele deve achar que é mais rentável jogar pelas regras, pois tais normas lhe favorecem com mais novas moedas do que todos os outros combinados, além de prejudicar o sistema e a validade de sua própria riqueza. (SATOSHI, 2008, p.4)

Entretanto assim como destacado por Taleb (2021), a problemática da falta de um valor intrínseco, e diversas diferenciações do bitcoin em relação a um investimento tradicional dado o fato que o criptoativo não realiza pagamentos de dividendos, possui grande imprevisibilidade sobre a sua rentabilidade futura. Além disso, não houve evidências que, de fato, a Blockchain provou ter alguma utilidade relevante e inovadora para a sociedade. Além disso, seu excessivo consumo de energia para um ativo que teria valor esperado não maior do que zero, fortalecem a corrente que acredita que os mineradores estariam desperdiçando recursos, naquilo que segundo Taleb (2021) seria um ativo extremamente frágil. Alguns desses pontos também são levantados pela Financial Times (2016), onde é questionado se a Blockchain não se trataria apenas de um “*hype*” (expressão idiomática em inglês utilizada para designar algo que está na moda, porém apresenta uma onda passageira).

Figura 2 - Consumo de energia utilizado ao longo dos anos na mineração de Bitcoins

Bitcoin Energy Consumption

Monthly, in terawatt-hours (TWh)



Fonte: Investopedia (2022)

Ao analisar a tabela acima, é possível de fato constatar o máximo consumo de energia utilizado na mineração do bitcoin, assim como é possível observar também como a atividade mineradora tem picos de alta e baixa significativos conforme a

cotação desse ativo digital sofre variações relevantes, reforçando a ideia que se caso o preço do bitcoin chegue a um valor irrisório, todos esses fatores de produção empregados na produção de novas moedas estarão permanentemente perdidos. Vale ressaltar que de acordo com a CNN (2021) países como a China proibiram a atividade mineradora, com o objetivo de assegurar a “ordem econômica e social do país”.

3.6 O QUE É O HALVING?

Como já descrito anteriormente, o Bitcoin tem uma oferta inicial limitada de 21 milhões de unidades de moeda, além disso, são emitidos e jogados na rede Bitcoin a cada dez minutos uma certa quantidade de moedas. O halving, também sendo conhecido por halvening, se trata de um evento que ocorre aproximadamente a cada 4 anos, no qual é estabelecido uma certa quantidade de bitcoin a serem criadas, que serão distribuídas em 210.000 blocos, sendo que tal quantidade cai pela metade.

Em 2009, eram emitidas cerca de 50 bitcoins a cada bloco (cerca de 10 minutos), com a ocorrência do primeiro halving no ano de 2012, passaram-se a ser emitidas 25 moedas a cada dez minutos, e assim sucessivamente.

É importante destacar também que as datas dos próximos halvings são conhecidas, pois ocorrem de forma programada e automática, não afetando as moedas que já estavam em circulação, e sim diminuindo a quantidade emitida de novas moedas.

Por fim, é importante destacar que o halving tem o papel de promover uma maior transparência no processo de criação do Bitcoin, explicitando o processo da emissão de novos Bitcoins, conferindo-lhe uma característica não inflacionária, visto que não são possíveis emissões além das previstas.

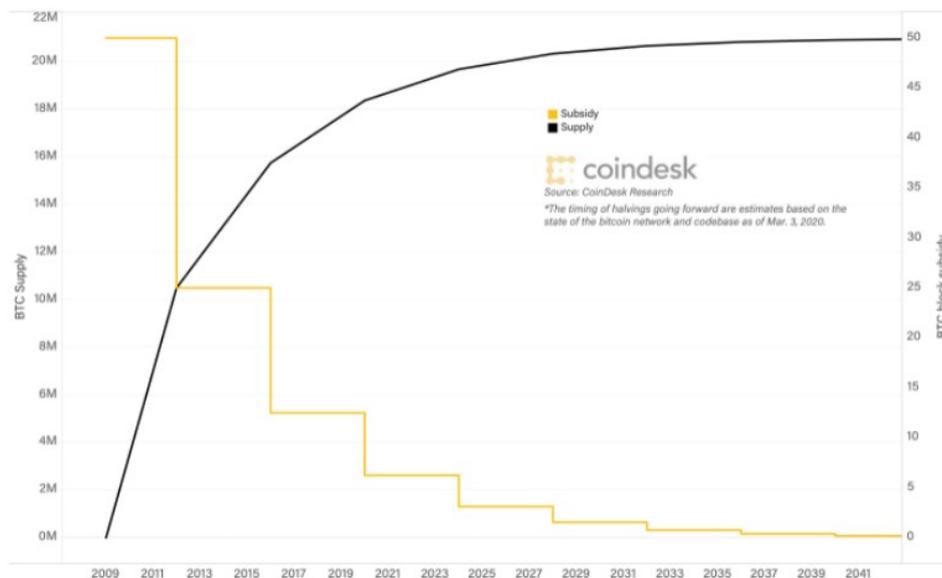
Ainda nessa seção, estarão presentes duas figuras nas quais ilustram as informações acima, a primeira delas ilustrando a quantidade de novas moedas criadas no período e como essa quantidade vai se reduzindo pela metade até tender a 0. Na segunda imagem, temos uma projeção em um plano cartesiano, na qual podemos ver com mais clareza como a oferta de moedas segue comportamento assintótico, assim como a emissão de novas moedas vão se reduzindo pela metade até tender a zero, evidenciando novamente, a característica de ser uma moeda com oferta limitada, e portando, apresentando comportamento deflacionário.

Figura 3 - Demonstração de como o halving interfere nos ganhos dos mineradores



Fonte: Mercado Bitcoin (2021)

Figura 4 - Expansão da base monetária do Bitcoin



Fonte: Coindesk (2021)

Vale destacar que o halving é uma das variáveis de maior importância para esta monografia, na qual investiga-se que ao reduzir a oferta de novas moedas, irá produzir um viés de alta no BTC. Ao analisar períodos anteriores essa prepositiva se mostrou verdadeira, contudo, não se pode creditar que eventos passados irão se repetir no futuro, sem evidências científicas.

4 METODOLOGIA

Neste trabalho, foi realizada uma competição entre modelos de classificação de saída binária com entradas multivariadas. Os modelos utilizados foram a regressão logística, o Gradient Boosting e o Random Forest. Como forma de comparação, utilizou-se a metodologia do SARIMAX, que é uma extensão do modelo ARIMA. Essa extensão permite levar em consideração a sazonalidade e incorporar variáveis exógenas na análise e previsão de séries temporais.

Além disso, foi realizada uma análise empírica com base em modelos heurísticos. Historicamente, observa-se que a cotação do Bitcoin tende a cair próximo à data do Halving e atingir suas máximas históricas cerca de um ano após esse evento. No entanto, é importante ressaltar que não há garantias de que essas tendências irão se manter nos próximos anos apenas com análises empíricas. Portanto, os modelos estatísticos possuem uma maior capacidade de realizar previsões mais robustas e fundamentadas.

Acerca do modelo de regressão logística, também conhecido como a inversa da LOGIT, se caracteriza pela possibilidade de poder converter o logaritmo de chances de um evento acontecer em probabilidade, dessa forma, obtendo a razão de possibilidades, que nos permite analisar com propriedade as chances de dado evento acontecer em relação as chances de não ocorrer. Por conta dessa característica, será possível avaliar com maior profundidade as possibilidades de ocorrência do Halving, dado, que se trata de um evento futuro que ainda não ocorreu.

A função logística apresenta como forma funcional o seguinte subscrito:

$$\text{logit}(p) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n$$

Onde: $\text{logit}(p)$ é a função logit inversa da probabilidade p

$\beta_0, \beta_1, \beta_2, \dots, \beta_n$ são os coeficientes que representam o efeito dos preditores (x_1, x_2, \dots, x_n) nas log-odds de p .

A forma funcional da regressão logística utilizada é dada por:

$$p = 1 / (1 + \exp(-z))$$

Onde: p é uma probabilidade e z é o valor da equação linear $\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n$.

No modelo elaborado, temos como variável de saída binária nomeada de “subiu ou caiu”, este representa a possibilidade de o preço do Bitcoin ter aumentado ou diminuído. A regressão logística é aplicada para estimar a probabilidade de 'Subiu_ou_caiu' ser verdadeiro ou falso com base nos padrões encontrados pelas variáveis de entrada e relações existentes nos dados históricos e, assim, prever a variável de saída (y) com base nessas informações. O objetivo é encontrar uma função de decisão que melhor se ajuste aos dados de treinamento e seja capaz de generalizar para novos dados não vistos durante o treinamento.

As variáveis de entrada, que compõem os nossos Betas são:

'Close_8MA': Representa a média móvel dos últimos 8 dias dos dados de fechamento. Essa variável captura a tendência geral do preço de fechamento e seu efeito nas probabilidades de classificação.

'Close_1', 'Close_2', 'Close_3': São defasagens (lags) de -1, -2 e -3 dias para a variável 'Close' que representa o fechamento do preço do BTC em determinado dia, ou seja, são os valores anteriores de fechamento. Essas variáveis capturam a relação entre os valores de fechamento passados e a probabilidade de classificação.

'Dummy1' e 'Dummy2': São variáveis que foram acrescentadas no modelo, para buscar capturar o efeito do Halving, nas datas dentro do intervalo de 6 meses anteriores e 6 meses subsequentes ao Halving, foi atribuído valor 1 e para demais datas, foi atribuído valor 0. A primeira Dummy captura o efeito do Halving de 2016 e a segunda captura o efeito do Halving de 2020.

O modelo de regressão logística não utiliza a estimação por mínimos quadrados ordinários, e sim a estimação por máxima verossimilhança que não possui fórmula funcional, porém tem como objetivo de estimar o nosso $x\beta$ máximo, que maximiza as chances de sucesso da nossa estimação, isto é, a ocorrência do halving (a principal variável explicativa deste modelo) e estar correlacionado com a alta do preço do bitcoin (a principal variável dependente deste modelo).

Para fins de comparação e competição entre os diferentes modelos, foram utilizadas as metodologias de aprendizado de máquina do Gradient Boosting e Random Forest. Que possuem as seguintes características:

O Gradient Boosting Classifier utiliza uma abordagem de conjunto de árvores de decisão para classificação. Ele é construído por meio de um processo iterativo, onde árvores de decisão mais simples são adicionadas sequencialmente para

melhorar o desempenho geral do modelo. A fórmula do Gradient Boosting Classifier é uma combinação ponderada das previsões individuais de cada árvore, onde as previsões de cada árvore são ajustadas de acordo com o erro cometido pelas previsões anteriores. A fórmula geralmente segue o seguinte padrão:

$$f(x) = b_0 + \alpha * h_1(x) + \alpha * h_2(x) + \dots + \alpha * h_n(x),$$

Onde $f(x)$ é a previsão final do modelo para uma determinada amostra de entrada x . b_0 é o viés (bias) do modelo, uma constante que representa uma previsão inicial. α é a taxa de aprendizado (learning rate), que controla a contribuição de cada árvore para a previsão final. $h_i(x)$ é a previsão da i -ésima árvore para a amostra de entrada x . O processo funciona da seguinte forma:

- Inicialmente, é criada uma árvore de decisão simples.
- Em cada iteração, uma nova árvore é adicionada ao conjunto para corrigir os erros das árvores anteriores.
- A nova árvore é ajustada aos resíduos (diferença entre as previsões atuais e os rótulos reais) das árvores anteriores.
- As previsões de todas as árvores são combinadas ponderadamente para obter a previsão final.

Dessa forma, o Gradient Boosting Classifier utiliza o princípio de "aprendizado por etapas" para construir um modelo forte a partir de várias árvores fracas, melhorando gradualmente sua capacidade de fazer previsões precisas.

Já o Random Forest Classifier, utiliza um conjunto de árvores de decisão para realizar classificação. Ele combina a técnica de bootstrap aggregating (ou "bagging") com a aleatoriedade na construção das árvores, o que o torna mais robusto e menos suscetível a overfitting. Ao contrário de um único modelo de árvore de decisão, o Random Forest Classifier utiliza várias árvores de decisão independentes para fazer previsões. Cada árvore é treinada em uma amostra aleatória do conjunto de dados original e, durante o processo de treinamento, cada árvore considera apenas um subconjunto aleatório das variáveis de entrada. A fórmula do Random Forest Classifier não é uma única equação como em alguns modelos lineares, em vez disso, a previsão do Random Forest Classifier é obtida através de uma combinação das previsões individuais de todas as árvores do conjunto. Para classificar uma amostra de entrada, cada árvore do conjunto emite uma votação, e a classe mais votada é selecionada

como a previsão final do modelo. A forma funcional do Random Forest Classifier é baseada em três princípios principais:

- Construção de árvores de decisão: Cada árvore é construída através da divisão recursiva do conjunto de dados em subconjuntos menores com base em características (variáveis de entrada) selecionadas aleatoriamente.

- Votação majoritária: Para fazer uma previsão, cada árvore vota em uma classe. A classe mais votada é selecionada como a previsão final do modelo.

- Redução do overfitting: A combinação de múltiplas árvores independentes e a aleatoriedade na seleção de características e amostras de treinamento ajuda a reduzir o overfitting e aumenta a generalização do modelo.

Dessa forma, o Random Forest Classifier utiliza a força coletiva de várias árvores de decisão para melhorar a precisão das previsões. Ao combinar a aleatoriedade na construção das árvores e o processo de votação, o modelo é capaz de lidar com dados complexos e realizar classificações precisas em uma variedade de problemas.

Por fim, foi utilizado para comparação a utilização do modelo autoprotelado do SARIMAX, uma variação do do modelo SARIMA, (*Seasonal autoregressive integrated moving average*) isto é modelo autorregressivo integrado de médias móveis sazonal de que corresponde a uma extensão do modelo ARIMA, muito utilizado para estimação e previsão de séries temporais não estacionárias. A saída deste modelo é uma variável quantitativa que reflete o preço do Bitcoin. A parte autorregressiva (AR) do modelo ARIMA indica que a variável de interesse é regressada em seus próprios valores defasados, já a parte de média móvel (MA) indica que o erro de regressão é na verdade uma combinação linear dos termos de erro, cujos valores ocorreram contemporaneamente e em vários momentos no passado. A parte integrada (I) indica que os valores de dados foram substituídos com a diferença entre seus valores e os valores anteriores e este processo diferenciador pode ter sido realizado mais de uma vez. A parte Sazonal (S) tem o objetivo de capturar um efeito sazonal do modelo, quando se há uma suspeita que o comportamento da série tende se repetir ao longo de certa quantidade de períodos padronizadas menores que 1 ano. Já a parte variável exógena (X) são informações externas que não são diretamente afetadas pela série temporal em si, mas podem influenciar seu comportamento. A metodologia do SARIMAX é bastante utilizada para séries temporais que apresentam comportamento

complexo e quando outras variáveis além da série temporal têm impacto significativo no seu comportamento.

4.1 Base de Dados

A base de dados é composta pela cotação diária do Bitcoin, disponível para acesso pelo Yahoo Finance, o período de recorte dos dados é de 17 de dezembro de 2014, até 17 de março de 2023. As informações sobre a base de dados diária utilizadas podem ser encontrada na API do Yahoo Finance disponibilizadas para utilização em python a partir do comando "import yfinance as yf" "ticker = yf.Ticker("BTC-USD)".

A divisão de treinamento e teste foi realizada de forma que 30% dos dados foram usados para teste e 70% para treinamento. Se encontrando da seguinte forma:

```
Treinamento: 2179 amostras - 69.9967876646322 %  
Teste: 934 amostras - 30.003212335367813 %
```

Fonte: Estimação própria

4.2 EVIDÊNCIAS EMPÍRICAS DA INFLUÊNCIA DO HALVING NAS VARIAÇÕES DOS PREÇOS DO BITCOIN

Inicialmente, é preciso destacar que por mais que exista a crença que o halving antecede períodos de alta dos preços do Bitcoin, não existem garantias de que isso, de fato, aconteça. Entretanto, com a redução da recompensa paga aos mineradores, caso não haja um aumento do preço do Bitcoin frente a outros ativos, muitos players podem acabar saindo desse mercado, dados os custos envolvidos no processo de mineração, que caso a cotação do bitcoin se encontre muito em baixa, torna o negócio pouco atrativo, e assim, torna a rede mais lenta, visto que, os pagamentos demoram a ser validados, pois há um número menor de máquinas trabalhando na autenticação das transações, todavia, com menos moedas sendo emitidas, o volume de transações vendedoras diminui, ocasionando um aumento dos preços do Bitcoin caso exista uma demanda compradora e assim permitindo a viabilidade financeira da atividade de mineração. De certo modo, o cenário descrito acima vem se repetindo durante todos os halvings até então. No ano de 2012, na qual ocorreu o primeiro halving, recompensava os mineradores por cada bloco encontrado em 50 BTC e era cotado

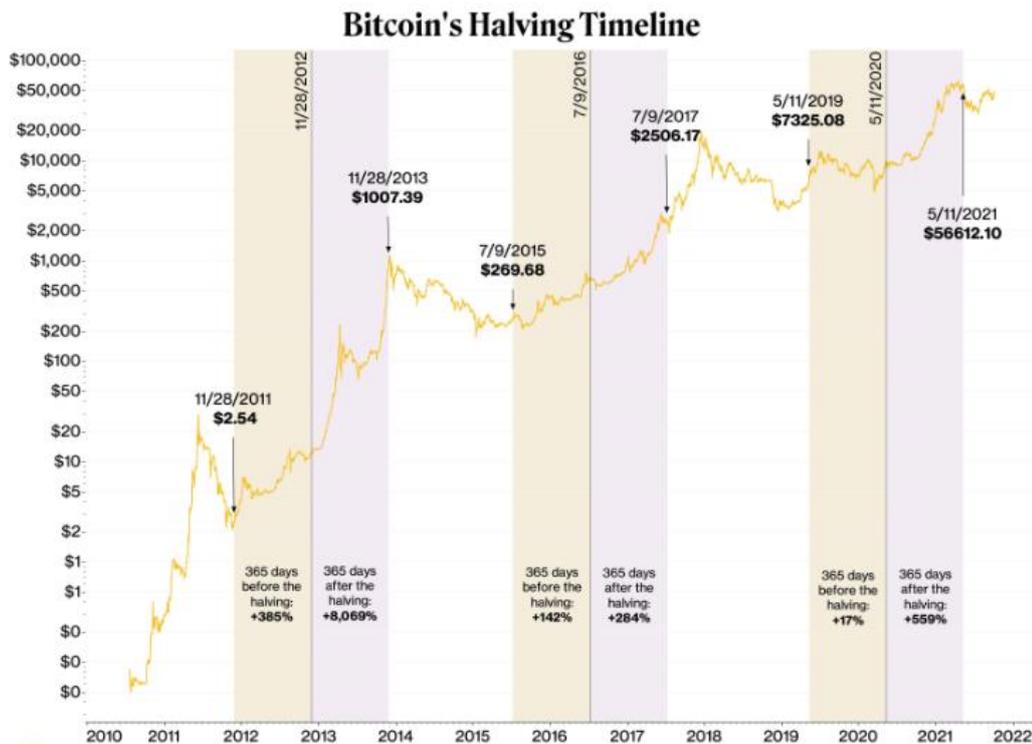
na casa dos 12 dólares, 125 após o halving, com a remuneração por bloco criptografado caindo para 25 BTC, a cotação do Bitcoin se encontrava na casa dos 127 dólares.

Em 9 de julho de 2016, o segundo halving apresentou comportamento semelhante, com a redução do pagamento de 25 BTC para 12,5 BTC por cada primeiro registro realizado na Blockchain, sua cotação saltou nos primeiros 150 dias após o halving de 650 dólares para mais de 758 dólares. Cerca de 365 dias após o segundo halving, o Bitcoin passou a ser cotado por mais de 2500 dólares.

No terceiro e mais recente halving, na qual há mais uma redução pela metade dos pagamentos em bitcoin aos minerados, há novamente um salto na sua cotação, saindo de cerca de 8800 dólares para quase 11000 dólares nos primeiros 150 dias. E após 1 ano do evento do terceiro halving, o Bitcoin se encontrava cotado em mais de 56000 dólares, uma alta de mais de 500%.

Da mesma forma, no ano anterior ao halving, a cotação do bitcoin se encontrava em baixa em relação a máxima atingida anteriormente.

Figura 5 - Gráfico em escala logarítmica da linha do tempo das cotações do Bitcoin em relação com o evento do Halving

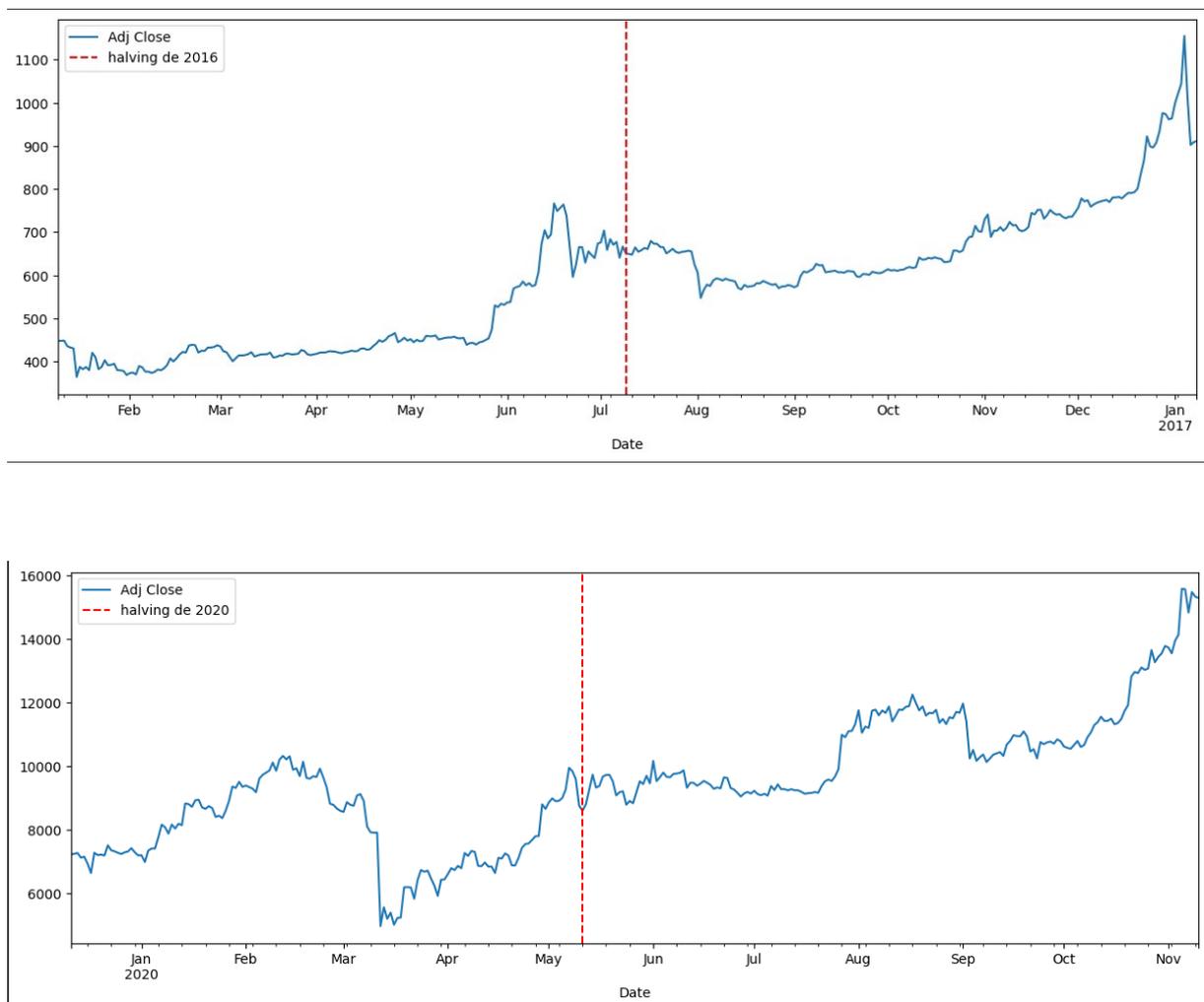


Fonte: Coindesk (2021)

a coloração em vermelho indica que o preço desse ativo se encontra em queda. Por fim as barras em azul são apenas marcadores das datas dos halvings.

Em última análise, ao realizar um recorte de um período de seis meses anteriores ao Halving e seis meses posteriores ao Halving, é possível identificar um padrão. Inicialmente, observa-se uma estagnação no crescimento da série, e após o evento, a série apresenta um aumento significativo, atingindo um patamar mais elevado. As figuras disponibilizadas a seguir são referentes aos Halvings que ocorreram nas datas de 9 de julho de 2016 e 11 de maio de 2020.

Figura 7 - Recortes dos Halvings anteriores dos anos de 2016 e 2020



Fonte: Estimação própria

5 – RESULTADOS

Ao gerar o relatório de classificação da competição de desempenho entre os modelos, obteve-se uma visão geral das métricas de desempenho do modelo, permitindo avaliar a capacidade do modelo de classificar corretamente se o preço do bitcoin subiu ou caiu. No modelo estimado, as principais métricas incluídas em um relatório de classificação de regressão logística são:

- Precisão (Precision): é a proporção de verdadeiros positivos (TP) em relação à soma dos verdadeiros positivos e falsos positivos (FP). Mede a capacidade do modelo de classificar corretamente os exemplos positivos.

- Recall (Revocação) ou Sensibilidade (Recall/Sensitivity): é a proporção de verdadeiros positivos (TP) em relação à soma dos verdadeiros positivos e falsos negativos (FN). Mede a capacidade do modelo de identificar corretamente os exemplos positivos.

- F1-Score: é a média harmônica entre a precisão e o recall. Fornece um equilíbrio entre essas duas métricas, sendo útil quando há um desequilíbrio entre as classes.

- Suporte (Support): é o número de ocorrências reais de cada classe no conjunto de dados.

- Weighted Average (Weight Avg): a métrica de média ponderada calcula a média ponderada das métricas de precisão, recall e F1-score para todas as classes, levando em consideração o suporte (número de ocorrências) de cada classe. Essa média ponderada atribui maior peso às classes com maior suporte, ou seja, com maior número de exemplos. É útil quando as classes têm distribuições desiguais no conjunto de dados.

- Macro Average (Macro Avg): a métrica de média aritmética calcula a média aritmética das métricas de precisão, recall e F1-score para todas as classes, independentemente do suporte de cada classe. Essa média trata todas as classes de forma igual, sem levar em consideração o desequilíbrio do número de exemplos entre as classes. É útil quando todas as classes têm a mesma importância.

Acurácia (Accuracy): a acurácia é uma métrica que calcula a proporção de exemplos corretamente classificados em relação ao total de exemplos. É uma medida geral do desempenho do modelo em classificar corretamente todas as classes. A acurácia não leva em consideração o desequilíbrio de classes, sendo mais adequada quando as classes têm distribuições balanceadas.

Sendo assim, pode-se notar que o modelo de regressão logística apresentou um desempenho relativamente superior aos seus concorrentes.

Figura 8 - Relatório de classificação dos modelos

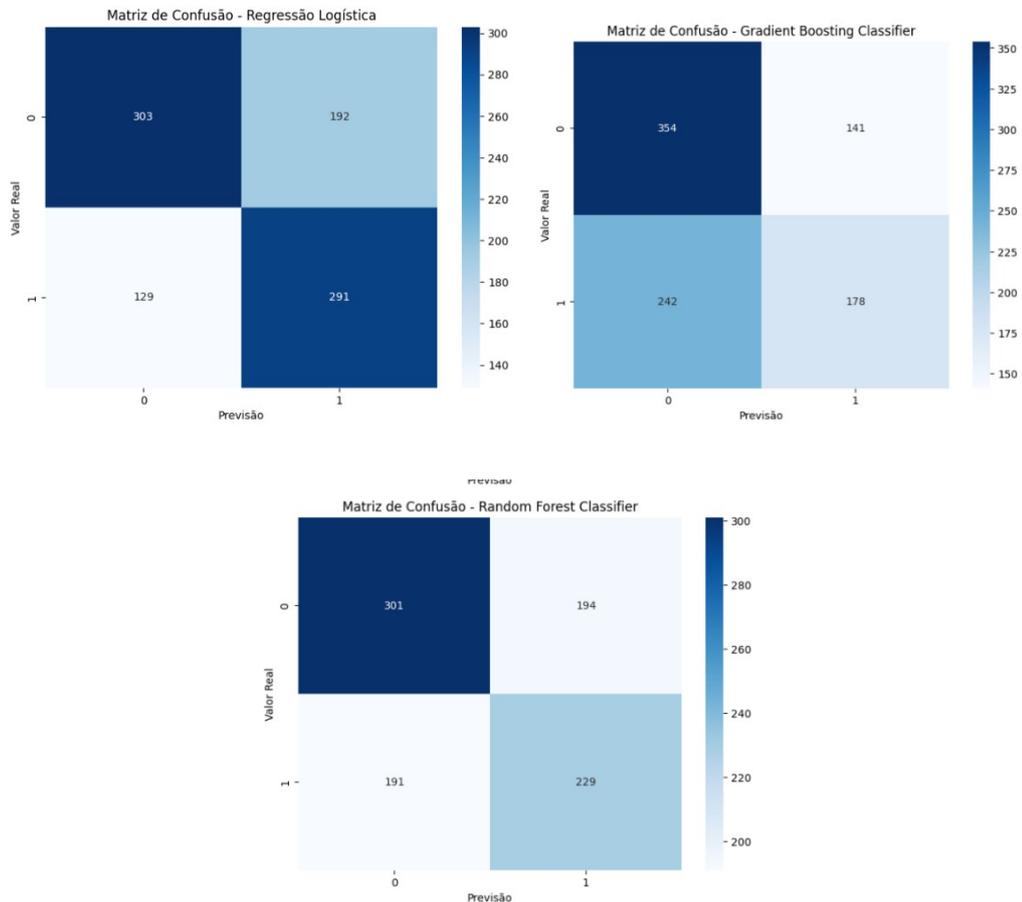
Logistic Regression Classifier:					
	precision	recall	f1-score	support	
0	0.67	0.68	0.67	491	
1	0.64	0.63	0.63	443	
accuracy			0.65	934	
macro avg	0.65	0.65	0.65	934	
weighted avg	0.65	0.65	0.65	934	
Gradient Boosting Classifier:					
	precision	recall	f1-score	support	
0	0.59	0.75	0.66	491	
1	0.60	0.42	0.49	443	
accuracy			0.59	934	
macro avg	0.59	0.58	0.57	934	
weighted avg	0.59	0.59	0.58	934	
Random Forest Classifier:					
	precision	recall	f1-score	support	
0	0.58	0.63	0.60	491	
1	0.55	0.50	0.52	443	
accuracy			0.57	934	
macro avg	0.56	0.56	0.56	934	
weighted avg	0.56	0.57	0.56	934	

Fonte: Estimação própria

Ao analisar a matriz de confusão, é possível obter uma compreensão mais precisa da qualidade dos ajustes desses modelos. Na matriz de confusão, os valores da diagonal principal representam as predições corretas, enquanto os falsos positivos e falsos negativos são representados pelos valores na diagonal secundária. No contexto da estimação do modelo utilizado, os valores "0" indicam que o preço do Bitcoin caiu, enquanto os valores "1" indicam que o preço aumentou. Mais uma vez, observa-se que o modelo de regressão logística apresenta um desempenho superior.

Embora o modelo Random Forest Classifier tenha acertado um número maior de valores TN (verdadeiros negativos), ele demonstrou maior dificuldade em relação aos valores TP (verdadeiros positivos).

Figura 9 - Matrizes de confusão dos diferentes modelos utilizados



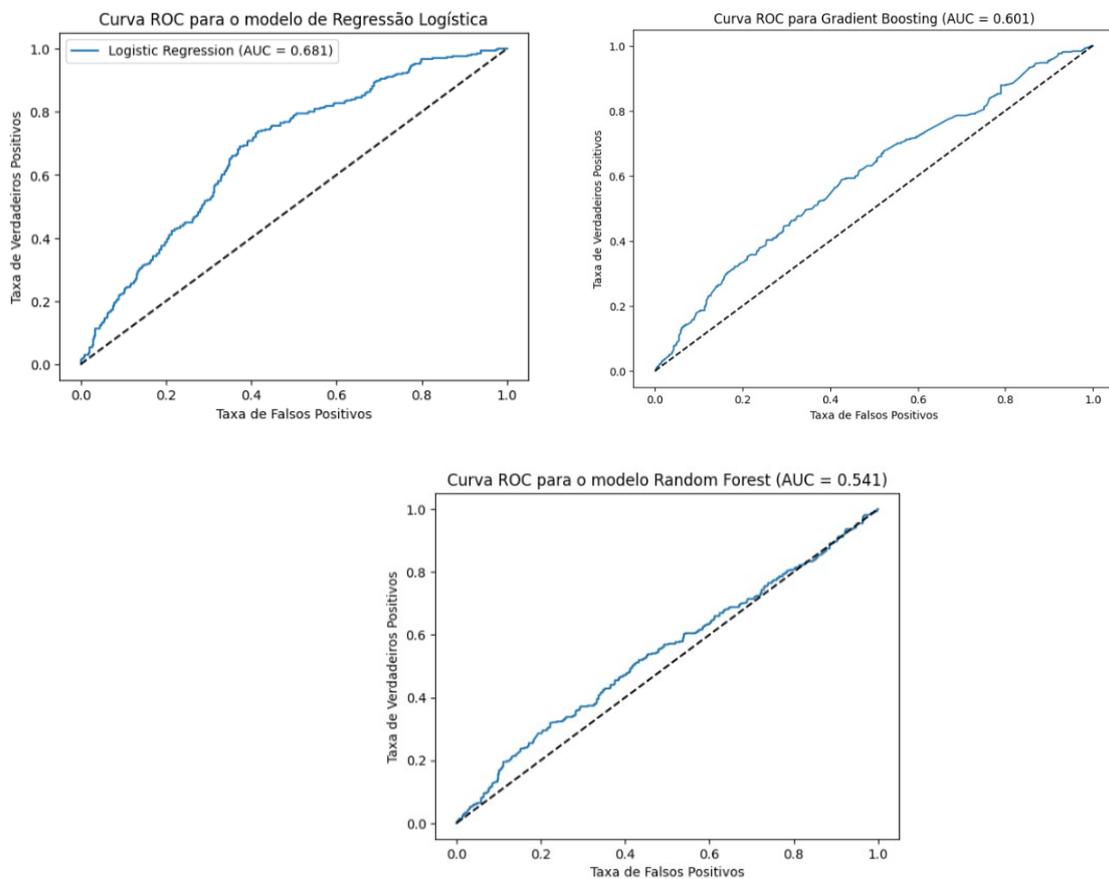
Fonte: Estimação própria

Por fim, utilizou-se o ferramental da Curva ROC (Curva Característica de Operação do Receptor), que é uma representação gráfica que ilustra o desempenho de um sistema classificador binário à medida que o limiar de discriminação varia. A curva ROC fornece uma medida abrangente da capacidade discriminativa do modelo, independentemente do ponto de corte escolhido. A área sob a curva ROC (AUC-ROC) é uma métrica comumente utilizada para comparar diferentes modelos. Quanto maior a AUC-ROC, melhor é a capacidade do modelo em distinguir entre classes positivas e negativas.

Além disso, mesmo diante de um desequilíbrio nas classes, como no caso do BTC, em que há mais exemplos de altas nos preços do que quedas, a curva ROC

pode fornecer uma avaliação mais robusta do desempenho do modelo. Após avaliar o desempenho nessas três métricas de avaliação, o modelo vencedor escolhido foi a Regressão logística. Esse modelo pode ser utilizado tanto para classificação quanto para previsão, e apresentou resultados suficientemente robustos para serem aplicados por interessados em estudar ou investir em Bitcoin.

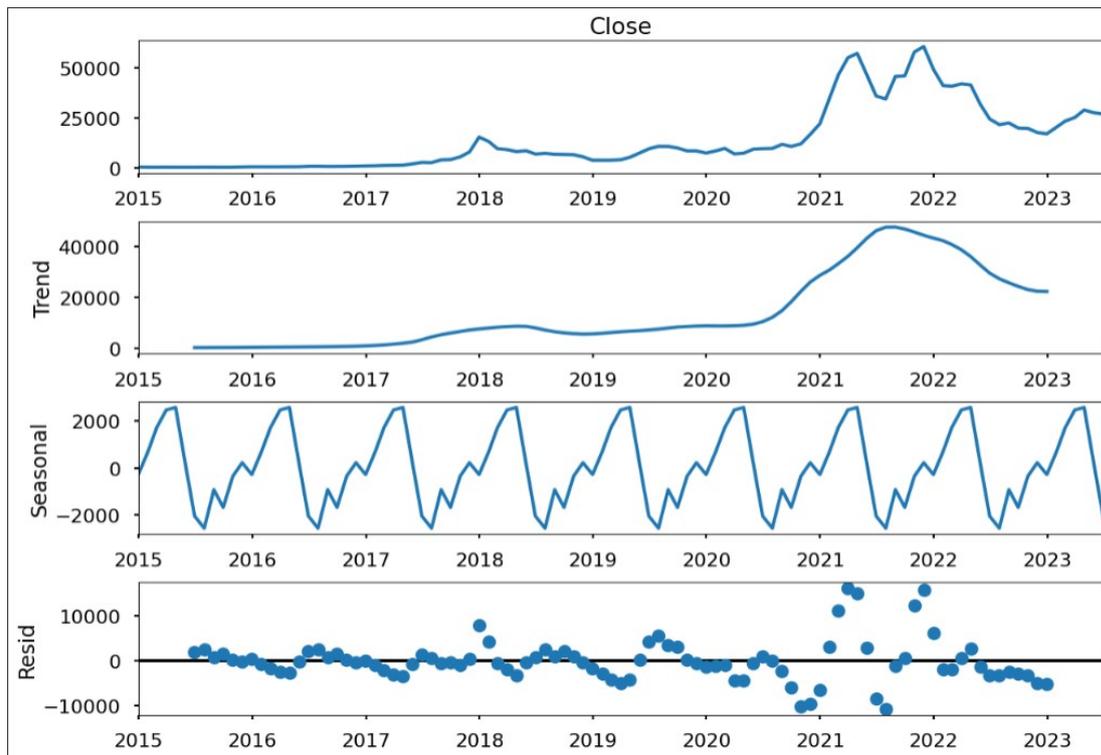
Figura 10 - Curvas ROC dos modelos elaborados



Fonte: Estimação própria

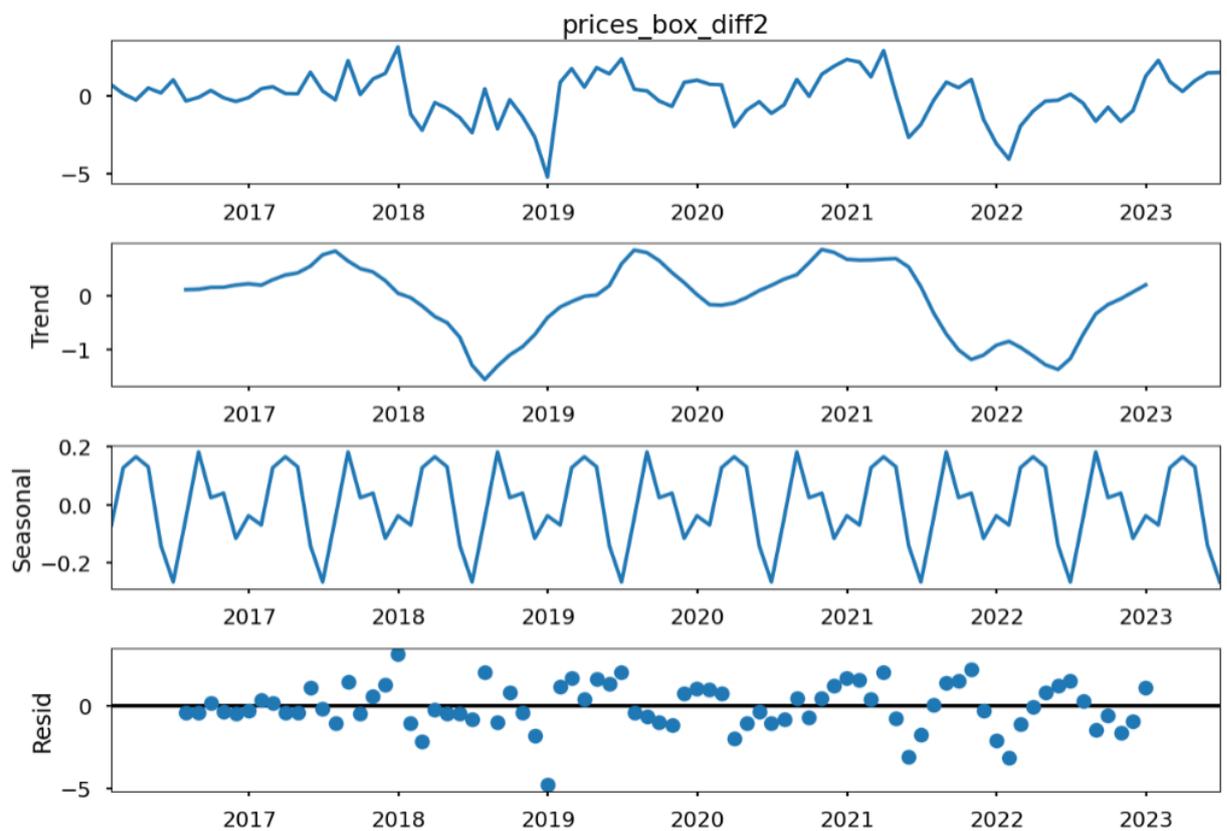
Já na estimação do modelo SARIMAX, utilizou-se dos valores das cotações mensais para elaborar um modelo de previsão univariada. Como de praxe, na metodologia ARIMA foi realizado o teste de Dickey-Fuller para verificar a estacionariedade da série, o seu nível de tendência e é realizada uma decomposição sazonal, com o objetivo de compreender os padrões dessa série. Vale ressaltar que nesse exemplo, dado a natureza explosiva dos dados é necessário realizar uma segunda diferenciação de Box-Cox para que os dados assumam características estacionárias.

Figura 11 – Primeira diferenciação da série temporal



Fonte: Estimação própria

Figura 12 – Segunda diferenciação da série temporal

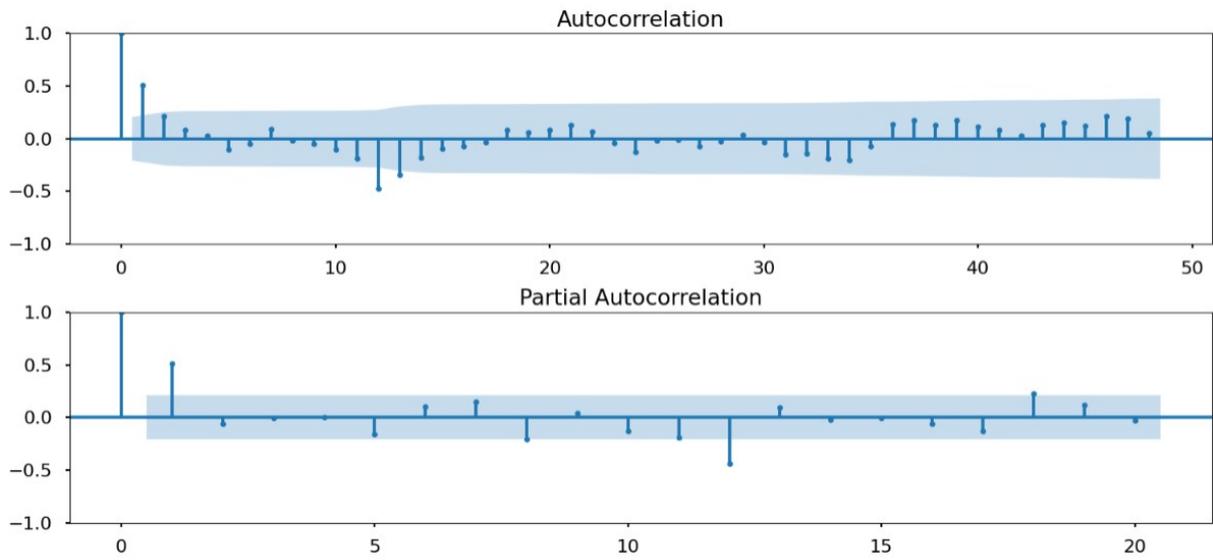


Fonte: Estimação própria

Após a segunda transformação de Box-Cox, é possível notar com os dados que a série se torna estacionária, isto é, uma série temporal estacionária é aquela cujas propriedades estatísticas, como a média, a variância e a auto correlação, são constantes ao longo do tempo, permitindo que seja possível continuar trabalhando com essa série, pois a metodologia de Box e Jenkins pressupõe que os parâmetros do modelo sejam estáveis ao longo do tempo. Em uma série não estacionária, os parâmetros podem variar significativamente, dificultando a modelagem adequada e a interpretação dos resultados.

Em sequência é realizado a análises da Função de Autocorrelação dos erros (FAC) e da Função de Autocorrelação Parcial dos Erros (FACP). A função de autocorrelação (FAC) mede a correlação entre os valores passados de uma série temporal e os valores atuais. Ela é calculada para diferentes defasagens (lags). A FAC é útil para identificar a presença de autocorrelação na série, ou seja, se há dependência entre as observações em diferentes momentos no tempo. Se a FAC mostrar autocorrelação significativa em certos lags, isso indica que a série possui uma estrutura temporal que pode ser explorada pelo modelo, Já a função de autocorrelação parcial (FACP) mede a correlação entre os valores passados da série, levando em consideração os efeitos indiretos causados por outras defasagens. A FACP ajuda a identificar a estrutura de dependência temporal específica que pode ser modelada pelo modelo ARIMA. Ela auxilia na determinação dos termos autorregressivos (AR) do modelo, fornecendo informações sobre quais lags são significativos para a predição da série. Vale ressaltar que só a análise gráfica, não é suficiente para ser definida a ordem e os parâmetros do modelo, para resolver esse gargalo será utilizado o Critério de Informação Akaike (AIC) que admite a existência de um modelo “real” que descreve os dados que é desconhecido, e tenta escolher dentre um grupo de modelos avaliados o que apresentar maior parcimônia, sendo assim o AIC leva em conta e penaliza a complexidade dos modelos e tende a favorecer a escolha de modelos mais simples.

Figura 13 - FAC e FACP



Fonte: Estimação própria

Figura 14 – Modelo Estimado

parameters	aic
22 (1, 0, 2, 0)	253.467860
23 (1, 0, 2, 1)	254.393020
19 (1, 0, 0, 1)	254.452580
46 (2, 1, 2, 0)	254.809256
40 (2, 0, 2, 0)	255.315369

SARIMAX Results

```

=====
Dep. Variable:          Weighted_Price_box      No. Observations:          102
Model:                 SARIMAX(1, 1, 0)x(2, 1, 0, 12)  Log Likelihood             -122.734
Date:                  Fri, 12 May 2023             AIC                        253.468
Time:                  23:43:26                     BIC                        263.422
Sample:                12-31-2014                       HQIC                       257.480
                    - 05-31-2023

```

Covariance Type: opg

	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.4506	0.087	5.201	0.000	0.281	0.620
ar.S.L12	-0.6539	0.077	-8.439	0.000	-0.806	-0.502
ar.S.L24	-0.4906	0.081	-6.084	0.000	-0.649	-0.333
sigma2	0.8307	0.107	7.729	0.000	0.620	1.041

```

=====
Ljung-Box (L1) (Q):          0.11      Jarque-Bera (JB):          2.21
Prob(Q):                     0.74      Prob(JB):                  0.33
Heteroskedasticity (H):     1.44      Skew:                      -0.00
Prob(H) (two-sided):        0.33      Kurtosis:                  3.77
=====

```

Fonte: Estimação própria

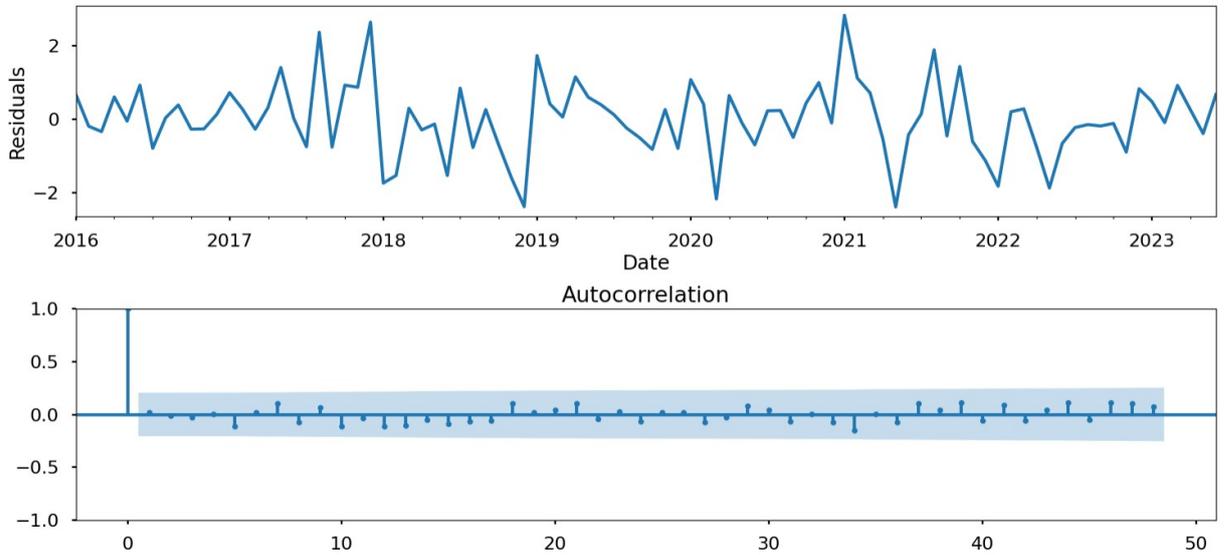
Com base, nesse ultima saída, classificamos o modelo SARIMAX (p,d,q)x(P,D,Q,S) estimado como um SARIMAX (1,1,0)x(2,1,0,12). Isto é, o parâmetro 'p' representa a ordem do componente autoregressivo (AR) não sazonal de ordem. O

parâmetro 'd' representa a ordem de diferenciação não sazonal de ordem 1 com transformação em log. Não há um parâmetro no componente de média móvel não sazonal. Na parte sazonal do modelo, O parâmetro 'P' representa a ordem do componente autoregressivo sazonal (SAR) de ordem 2, o parâmetro 'D' representa a ordem de diferenciação sazonal, nesse caso, há novamente 1 diferenciação com transformação logarítmica. O parâmetro 'Q' que representa a ordem do componente de média móvel sazonal (SMA) possui ordem 0. O parâmetro 'S' especifica o número de observações em um ciclo sazonal, como foi utilizada uma base de dados mensal, seu valor é igual a 12. Dessa forma, o modelo SARIMAX(1, 1, 0)x(2, 1, 0, 12) é um modelo SARIMA que incorpora um componente autoregressivo e um componente de média móvel, tanto para os componentes não sazonais quanto para os componentes sazonais, permitindo incluir variáveis exógenas para uma melhor explicação do modelo.

Após encontrar o modelo que deverá ser estimado, é realizada a decomposição STL (Seasonal and Trend decomposition using Loess) dos resíduos do modelo SARIMAX(1, 1, 0)x(2, 1, 0, 12) afim de garantir que os resíduos do modelo não possuem algum padrão não capturado pela estrutura do modelo. Se os resíduos apresentarem padrões significativos ou não forem estacionários, isso pode indicar que o modelo precisa ser ajustado ou que outros fatores precisam ser considerados na modelagem. Entretanto é possível notar um comportamento estacionário dos resíduos e erros não correlatados, o que indica um bom ajuste do nosso modelo, mesmo sem a inclusão de uma variável exógena.

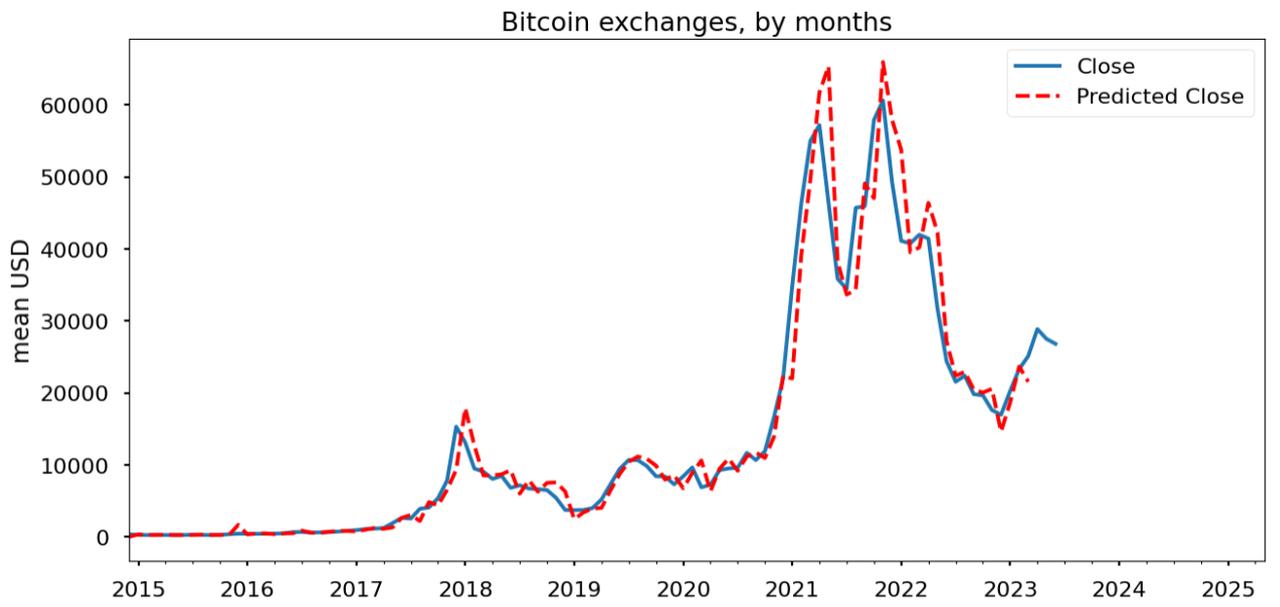
A partir disso, é realizada a transformação inversa de Box-Cox com o objetivo de retornar a série transformada para sua escala original, visando realizar previsões para períodos subsequentes. Inicialmente, foi testado o ajuste do modelo comparando os valores reais com os valores previstos. Observou-se um excelente ajuste, mesmo considerando o perfil explosivo da série. Dessa forma, ao realizar a previsão para o ano de 2024, o modelo indicou uma possibilidade de alta logo após o período do próximo halving. Essa é uma indicação relevante, considerando as hipóteses anteriores que sugerem que o halving antecede algumas das maiores altas históricas do Bitcoin.

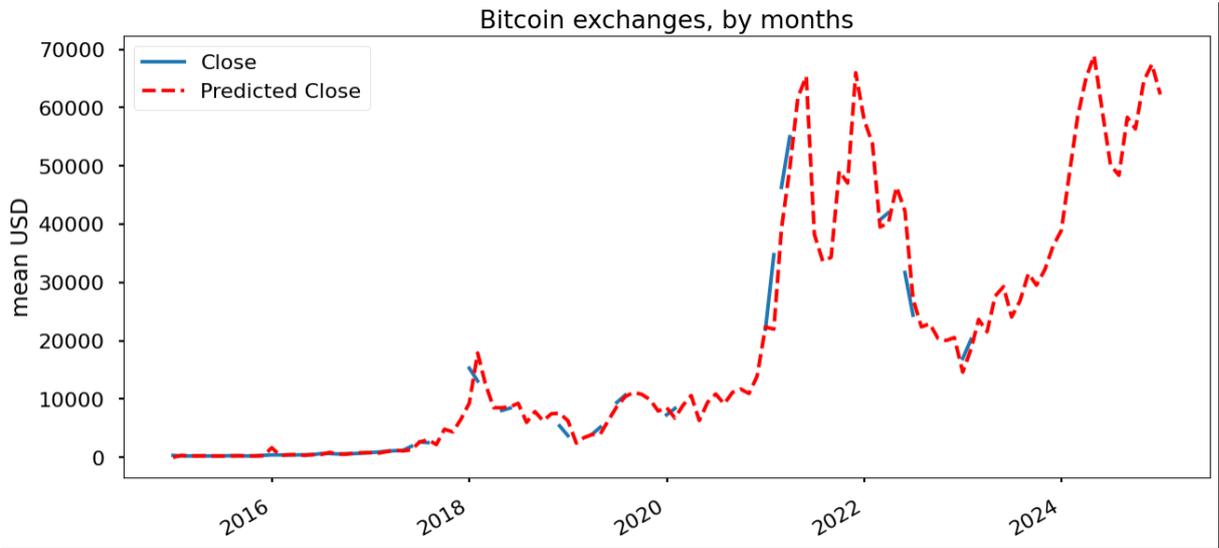
Figura 15 - – Decomposição STL



Fonte: Estimação própria

Figura 16 - – Previsão do modelo SARIMAX





Fonte: Estimação própria

6 – Considerações finais

Considerando os aspectos discutidos nos capítulos anteriores, cujo principal objetivo foi investigar se o halving desempenha um papel determinante nas altas dos preços da criptomoeda Bitcoin, destacam-se algumas descobertas que merecem ênfase. Entre elas, pode-se observar a forte relação entre as variáveis estudadas. Com base nas análises empíricas realizadas, é evidente a existência de uma correlação significativa entre essas variáveis, o que indica a possibilidade de uma alta do Bitcoin em um horizonte próximo. Essas conclusões respaldam a hipótese inicialmente proposta, reforçando a importância do fenômeno do halving como um impulsionador dos preços da criptomoeda. Ao considerar a natureza do mercado de criptomoedas e a relevância histórica do halving, pode-se inferir que esse evento tem o potencial de desencadear um aumento significativo no valor do Bitcoin.

No contexto da presente monografia, ao avaliar a competição entre os métodos de classificação, observa-se a eficácia da regressão logística em relação a outros métodos mais avançados de aprendizado de máquina. A regressão logística se destaca ao proporcionar uma classificação mais precisa tanto para os períodos em que o preço do ativo digital aumenta, quanto para os períodos de queda nos preços. Esses resultados evidenciam o potencial dessa abordagem como uma ferramenta valiosa para a previsão e classificação de movimentos do mercado.

Outrossim, um aspecto de destaque reside no desempenho do modelo SARIMAX, que diferentes dos outros métodos, possui como característica o fato de ser univariado e autoprojetivo. A projeção resultante indica uma tendência altista significativa para os períodos futuros, especialmente em relação ao próximo halving previsto para 2024. Essa constatação está em consonância com as evidências empíricas apresentadas nesta monografia.

É importante ressaltar, no entanto, que nenhum modelo é infalível. Apesar dos excelentes resultados obtidos, a natureza explosiva dos dados e a volatilidade do mercado representam desafios significativos para qualquer metodologia de previsão e classificação. Portanto, embora a regressão logística tenha se mostrado promissora e eficiente, é necessário considerar as limitações inerentes à previsão de movimentos de preços em um ambiente tão dinâmico e complexo como o do Bitcoin.

É importante ressaltar também, que um horizonte de previsão de cerca de um ano é considerado relativamente longo e está sujeito a volatilidades de curto e longo prazo, que podem diferenciar o valor projetado do valor real. Para realizar projeções em horizontes de previsão mais extensos, é necessário contar com uma grande quantidade de dados históricos, levando em consideração que a cotação do BTC teve início em 2014. Dessa forma, quanto mais próximo do halving, mais consistentes serão as previsões. Além disso, embora as evidências apontem para a possibilidade de uma alta no Bitcoin após o próximo halving, é necessário continuar monitorando e avaliando os desenvolvimentos do mercado e os fatores externos que podem influenciar os preços das criptomoedas.

Contudo, apesar das limitações existentes, vale destacar algumas contribuições que este estudo proporcionou. Além de apresentar modelos de previsão e classificação, a pesquisa permitiu um maior entendimento do Bitcoin como um todo, abordando os motivos que levaram à sua criação por Satoshi Nakamoto, os mecanismos pelos quais opera, suas principais características e filosofias, bem como sua diferenciação em relação às moedas fiduciárias. Além disso, o estudo trouxe à tona questões fundamentais relacionadas à natureza descentralizada e à tecnologia subjacente ao Bitcoin, como a blockchain. Ao investigar esses aspectos, foi possível compreender o potencial disruptivo do Bitcoin no sistema financeiro e explorar suas possíveis aplicações em diversos setores da economia.

Vale ressaltar que novos estudos que aprofundem a questão das limitações dos modelos preditivos são extremamente bem-vindos. O crescimento da inteligência artificial e dos modelos avançados de aprendizado de máquina permitem a utilização de modelos cada vez mais complexos, porém se encontram fora do escopo desse trabalho. Também está aberto um espaço para estudos que avaliem o comportamento de bolha no bitcoin, ou que avaliem se a cotação reflete o seu valor real.

Por fim, este trabalho finaliza compreendendo o seu papel disruptivo na área, visto que, os estudos acadêmicos na área de economia sobre ativos digitais ainda são bastante limitados, tanto no Brasil como no mundo, este trabalho buscou compreender os mecanismos de funcionamento da criptomoeda que o fazem ter uma característica de uma moeda deflacionária, investigando o papel do halving na determinação do preço do bitcoin.

REFERÊNCIAS

Autor de livros como "Cisne Negro" e "Antifragil", Nassim Taleb chama bitcoin de "doença contagiosa", Valor Econômico, 2022. Disponível em <<https://valor.globo.com/financas/noticia/2022/01/18/autor-de-livros-como-cisne-negro-e-antifragil-nassim-taleb-chama-bitcoin-de-doenca-contagiosa.html>> Acesso em 17 de janeiro de 2022.

BAUR, Aaron W., et al. "Cryptocurrencies as a disruption? empirical findings on user adoption and future potential of bitcoin and co." *Conference on e-Business, e-Services and e-Society*. Springer, Cham, 2015.

Bitcoin and blockchain: the future of money or just hype? Financial Times, 2016. Disponível em <<https://www.ft.com/content/3bea303c-7a7e-11e6-b837-eb4b4333ee43>> Acesso em 20 de janeiro de 2022.

Bitcoin clock, Buybitcoinworldwide, 2022. Disponível em <<https://www.buybitcoinworldwide.com/bitcoin-clock>> Acesso em 20 de janeiro de 2022.

Bitcoin Halving, Explained, Coindesk, 2021. Disponível em <<https://www.coindesk.com/learn/2020/03/24/bitcoin-halving-explained>> Acesso em 19 de janeiro de 2022.

Bitcoin Pizza Day: a história da refeição mais cara de todos os tempos, Exame, 2021. Disponível em <<https://exame.com/future-of-money/criptoativos/bitcoin-pizza-day-a-historia-da-refeicao-mais-cara-de-todos-os-tempos>> Acesso em 18 de janeiro de 2022.

BLANCHARD, Olivier et al Macroeconomia, Editora Person, São Paulo, 2011

CAMACHO, Tatiana Silveira; DA SILVA, Guilherme Jonas Costa. Criptoativos: Uma Análise do Comportamento e da Formação do Preço do Bitcoin. *Revista de Economia*, v. 39, n. 68, 2019.

China proíbe mineração e declara ilegais transações com criptomoedas no país, CNN, 2021. Disponível em <<https://www.cnnbrasil.com.br/business/china-amplia-restricoes-e-proibe-mineracao-de-criptomoedas-em-todo-o-pais/>> Acesso em 10 de agosto de 2022.

CORAZZA, Gentil. Os Bancos Centrais e sua ambivalência público-privada. *Nova Economia*, v. 11, n. 1, 2001.

FERGUSON, Niall. A ascensão do dinheiro: a história financeira do mundo. São Paulo, Editora Planeta, 2009

HAYEK, Friedrich A. Desestatização do dinheiro. LVM Editora, 2017.

How does Bitcoin mining work, Investopedia, 2022. Disponível em <<https://www.investopedia.com/tech/how-does-bitcoin-mining-work>> Acesso em 20 de janeiro de 2022.

- IORIO, Ubiratan. Cantillon, os ciclos econômicos e a não-neutralidade da moeda. Publicado pelo Instituto Ludwig von Mises em, v. 15.
- MANKIWI, N. Gregory et al. Macroeconomics. New York: Worth Publishers, 2014.
- MENGER, Carl, A origem do Dinheiro, São Paulo Ludwig von Mises Institute, 2011
- Monte dei Paschi di Siena, 2017. Disponível em <<https://www.gruppompis.it/en/about-us/history.html>> Acesso em 10 de agosto de 2022.
- NAKAMOTO, Satoshi; BITCOIN, A. A peer-to-peer electronic cash system. Bitcoin. – URL: <https://bitcoin.org/bitcoin.pdf>, v. 4, 2008.
- NICHOLAS TALEB, Nassim. Bitcoin, currencies, and fragility. **Quantitative Finance**, v. 21, n. 8, p. 1249-1255, 2021.
- NORDHAUS, Samuelson. Economia, editora AMGH; 19ª edição, 2012.
- O que é o Halving e Como Funciona? Tudo o Que Você precisa saber! Mercado Bitcoin 2021. Disponível em < <https://blog.mercadobitcoin.com.br/o-que-e--halving-no-bitcoin>> Acesso em 19 de janeiro de 2022.
- REISMAN, George. Deflação, prosperidade e padrão-ouro, 2010.
- Sveriges Riksbank, 2019. Disponível em <<https://www.riksbank.se/en-gb/about-the-riksbank/history/the-riksbanks-350th-anniversary/>> Acesso em 10 de agosto de 2022.
- ULRICH, Fernando. Bitcoin: a moeda na era digital. LVM Editora, 2017.
- YERMACK, David. Is Bitcoin a real currency? An economic appraisal. In: Handbook of digital currency. Academic Press, 2015.

APÊNDICE A – Código em Python das regressões

```

import datetime
import numpy as np
import pandas as pd
import seaborn as sns
import pandas_datareader.data as web

from sklearn import (
    linear_model, metrics, neural_network, pipeline, preprocessing,
    model_selection
)

import matplotlib.pyplot as plt
%matplotlib inline
from datetime import datetime

!pip install yfinance --upgrade --no-cache-dir
import yfinance as yf

start = datetime(2014,12,9)
end = datetime.now().date().isoformat()
symbol = "BTC-USD"

df = yf.download(symbol, start=start, end=end)
df.tail(2)

df['Adj Close'].plot(figsize=(15,5))
df = df.drop(columns=['Volume'])
df.plot(figsize=(15,5))

#recortes dos halvings anteriores - 6 meses
## 9 de julho de 2016

start = datetime(2016,1,9)
end = datetime(2017,1,9) #datetime.now().date().isoformat()
symbol = "BTC-USD"
df = yf.download(symbol, start=start, end=end)

start = datetime(2016, 1, 9)
end = datetime(2017, 1, 9)
symbol = "BTC-USD"
df = yf.download(symbol, start=start, end=end)

plt.figure(figsize=(15, 5))
df['Adj Close'].plot()

```

```

plt.axvline(x=datetime(2016, 7, 9), color='red', linestyle='--',
label='halving de 2016')
plt.legend()
plt.show()

## Halving de 2020 - 11 de maio de 2020 - 6 meses
start = datetime(2019,12,11)
end = datetime(2020,11,11) #datetime.now().date().isoformat()
symbol = "BTC-USD"
df = yf.download(symbol, start=start, end=end)
df['Adj Close'].plot(figsize=(15,5))

start = datetime(2019, 12, 11)
end = datetime(2020, 11, 11)
symbol = "BTC-USD"
df = yf.download(symbol, start=start, end=end)

plt.figure(figsize=(15, 5))
df['Adj Close'].plot()
plt.axvline(x=datetime(2020, 5, 11), color='red', linestyle='--',
label='halving de 2020')
plt.legend()
plt.show()

from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report

start = datetime(2014,12,9)
end = datetime.now().date().isoformat()
symbol = "BTC-USD"
df = yf.download(symbol, start=start, end=end)

# Remover a coluna 'Adj Close'
df.drop('Adj Close', axis=1, inplace=True)

# Criar variáveis dummy para os períodos entre 9 de janeiro de 2016 e 9
de Janeiro de 2017 e entre 11 de novembro de 2019 e 11 de novembro de
2020
df['Dummy1'] = np.where((df.index >= '2016-01-09') & (df.index < '2017-
01-09'), 1, 0)
df['Dummy2'] = np.where((df.index >= '2019-11-11') & (df.index < '2020-
11-11'), 1, 0)

df['Subiu_ou_caiu'] = np.where(df['Close'].shift(-1) > df['Close'],
'0', '1')
df.dropna(inplace=True) # Remover as linhas com valores ausentes
resultantes dos deslocamentos

```

```

# Adicionar deslocamentos de -1, -2 e -3 para a coluna 'Close'
df['Close_1'] = df['Close'].shift(-1)
df['Close_2'] = df['Close'].shift(-2)
df['Close_3'] = df['Close'].shift(-3)

df.dropna(inplace=True) # Remover as linhas com valores ausentes
resultantes dos deslocamentos

# Adicionar uma coluna para a média móvel dos últimos 8 dias dos dados
de fechamento
df['Close_8MA'] = df['Close'].rolling(window=8).mean()

# Remover as linhas com valores ausentes resultantes da média móvel
df.dropna(inplace=True)

# Separar os dados em recursos (X) e rótulos (y)
X = df[['Close_8MA', 'Close_1', 'Close_2', 'Close_3', 'Dummy1',
'Dummy2']]
y = df['Subiu_ou_caiu']

# Dividir os dados em conjuntos de treinamento e teste
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.30, random_state=200)

# Treinar o modelo de regressão logística
logmodel = LogisticRegression()
logmodel.fit(X_train,y_train)

# Realizar as previsões para o conjunto de teste
predictions = logmodel.predict(X_test)

# Imprimir o relatório de classificação
print(classification_report(y_test,predictions))

from sklearn.ensemble import GradientBoostingClassifier
from sklearn.ensemble import RandomForestClassifier from
sklearn.svm import SVC

# Treinando um modelo Gradient Boosting
gb_clf = GradientBoostingClassifier(random_state=42)
gb_clf.fit(X_train, y_train)

# Fazendo previsões com o modelo Gradient Boosting
y_pred_gb = gb_clf.predict(X_test)

# Gerando um relatório de classificação para o modelo Gradient Boosting
print("Gradient Boosting Classifier:\n", classification_report(y_test,
y_pred_gb))

```

```

# Treinando um modelo Random Forest
rf_clf = RandomForestClassifier(random_state=42)
rf_clf.fit(X_train, y_train)

# Fazendo previsões com o modelo Random Forest
y_pred_rf = rf_clf.predict(X_test)

# Gerando um relatório de classificação para o modelo Random Forest
print("Random Forest Classifier:\n", classification_report(y_test,
y_pred_rf))

from sklearn.metrics import confusion_matrix
import matplotlib.pyplot as plt
import seaborn as sns

# Matriz de confusão para Regressão Logística
cm_logmodel = confusion_matrix(y_test, predictions)

# Plotar a matriz de confusão para Regressão Logística
plt.figure(figsize=(8, 6))
sns.heatmap(cm_logmodel, annot=True, fmt="d", cmap="Blues")
plt.title("Matriz de Confusão - Regressão Logística")
plt.xlabel("Previsão")
plt.ylabel("Valor Real")
plt.show()

# Matriz de confusão para Gradient Boosting Classifier
cm_gb = confusion_matrix(y_test, y_pred_gb)

# Plotar a matriz de confusão para Gradient Boosting Classifier
plt.figure(figsize=(8, 6))
sns.heatmap(cm_gb, annot=True, fmt="d", cmap="Blues")
plt.title("Matriz de Confusão - Gradient Boosting Classifier")
plt.xlabel("Previsão")
plt.ylabel("Valor Real")
plt.show()

# Matriz de confusão para Random Forest Classifier
cm_rf = confusion_matrix(y_test, y_pred_rf)

# Plotar a matriz de confusão para Random Forest Classifier
plt.figure(figsize=(8, 6))
sns.heatmap(cm_rf, annot=True, fmt="d", cmap="Blues")
plt.title("Matriz de Confusão - Random Forest Classifier")
plt.xlabel("Previsão")
plt.ylabel("Valor Real")
plt.show()
from sklearn.metrics import roc_curve, roc_auc_score

```

```

from sklearn.preprocessing import LabelEncoder
import matplotlib.pyplot as plt

# Converter as classes em valores numéricos
le = LabelEncoder()
y_test = le.fit_transform(y_test)

# Calcular as taxas de verdadeiros positivos e falsos positivos para o
modelo de Regressão Logística
y_pred_proba = logmodel.predict_proba(X_test)[:, 1]
fpr, tpr, _ = roc_curve(y_test, y_pred_proba)
auc_score = roc_auc_score(y_test, y_pred_proba)

# Plotar a curva ROC para o modelo de Regressão Logística
plt.plot(fpr, tpr, label=f'Logistic Regression (AUC =
{auc_score:.3f})')
plt.plot([0, 1], [0, 1], 'k--')
plt.xlabel('Taxa de Falsos Positivos')
plt.ylabel('Taxa de Verdadeiros Positivos')
plt.title('Curva ROC para o modelo de Regressão Logística')
plt.legend()
plt.show()

# Calcular as taxas de verdadeiros positivos e falsos positivos para o
modelo Gradient Boosting
y_pred_proba = gb_clf.predict_proba(X_test)[:, 1]
fpr, tpr, _ = roc_curve(y_test, y_pred_proba)
auc_score = roc_auc_score(y_test, y_pred_proba)

# Plotar a curva ROC para o modelo Gradient Boosting
fig, ax = plt.subplots(figsize=(8, 6))
ax.plot(fpr, tpr)
ax.plot([0, 1], [0, 1], 'k--')
ax.set_xlabel('Taxa de Falsos Positivos')
ax.set_ylabel('Taxa de Verdadeiros Positivos')
ax.set_title(f'Curva ROC para Gradient Boosting (AUC =
{auc_score:.3f})')
plt.show()

# Treinar o modelo Random Forest
rf_clf = RandomForestClassifier(n_estimators=100, max_depth=5,
random_state=42)
rf_clf.fit(X_train, y_train)

# Calcular as taxas de verdadeiros positivos e falsos positivos para o
modelo Random Forest
y_pred_proba = rf_clf.predict_proba(X_test)[:, 1]
fpr, tpr, _ = roc_curve(y_test, y_pred_proba)
auc_score = roc_auc_score(y_test, y_pred_proba)

```

```
# Plotar a curva ROC para o modelo Random Forest
plt.plot(fpr, tpr)
plt.plot([0, 1], [0, 1], 'k--')
plt.xlabel('Taxa de Falsos Positivos')
plt.ylabel('Taxa de Verdadeiros Positivos')
plt.title(f'Curva ROC para o modelo Random Forest (AUC =
{auc_score:.3f}) ')
plt.show()
```

APÊNDICE B – Modelo de previsão SARIMAX

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib as mpl
from scipy import stats
import statsmodels.api as sm
import warnings
from itertools import product
from datetime import datetime
warnings.filterwarnings('ignore')
plt.style.use('seaborn-poster')
import yfinance as yf

start = datetime(2014,12,9)
end = datetime.now().date().isoformat()
symbol = "BTC-USD"

df = yf.download(symbol, start=start, end=end)

from datetime import datetime

start = datetime(2014,12,9)
end = datetime.now().date().isoformat()
symbol = "BTC-USD"

df = yf.download(symbol, start=start, end=end)
df['Date'] = pd.to_datetime(df.index, unit='s')

# Resampling to daily frequency
df = df.resample('D').mean()

# Resampling to monthly frequency
df_month = df.resample('M').mean()

# Resampling to annual frequency
df_year = df.resample('A-DEC').mean()

# Resampling to quarterly frequency
df_Q = df.resample('Q-DEC').mean()

# PLOTS
fig = plt.figure(figsize=[15, 7])
plt.suptitle('Bitcoin exchanges, mean USD', fontsize=22)

plt.subplot(221)
```

```

plt.plot(df.Close, '-', label='By Days')
plt.legend()

plt.subplot(222)
plt.plot(df_month['Close'], '-', label='By Months')

plt.subplot(223)
plt.plot(df_month['Close'], '-', label='By Quarters')
plt.legend()

plt.subplot(224)
plt.plot(df_month['Close'], '-', label='By Years')
plt.legend()

# plt.tight_layout()
plt.show()

plt.figure(figsize=[15,7])
sm.tsa.seasonal_decompose(df_month['Close']).plot()
print("Dickey-Fuller test: p=%f" %
sm.tsa.stattools.adfuller(df_month['Close'])[1])
plt.show()

df_month['Weighted_Price_box'], lambda = stats.boxcox(df_month['Close'])
print("Dickey-Fuller test: p=%f" %
sm.tsa.stattools.adfuller(df_month['Close'])[1])

df_month['prices_box_diff'] = df_month['Weighted_Price_box'] -
df_month['Weighted_Price_box'].shift(12)
print("Dickey-Fuller test: p=%f" %
sm.tsa.stattools.adfuller(df_month['prices_box_diff'])[12:])[1])

# Regular differentiation
df_month['prices_box_diff2'] = df_month.prices_box_diff -
df_month.prices_box_diff.shift(1)
plt.figure(figsize=(15,7))

# STL-decomposition
sm.tsa.seasonal_decompose(df_month.prices_box_diff2[13:]).plot()
print("Dickey-Fuller test: p=%f" %
sm.tsa.stattools.adfuller(df_month.prices_box_diff2[13:])[1])
plt.show()

plt.figure(figsize=(15,7))
ax = plt.subplot(211)
sm.graphics.tsa.plot_acf(df_month.prices_box_diff2[13:].values.squeeze(
), lags=48, ax=ax)
ax = plt.subplot(212)

```

```

sm.graphics.tsa.plot_pacf(df_month.prices_box_diff2[13:].values.squeeze
(), lags=20, ax=ax)
plt.tight_layout()
plt.show()

# Initial approximation of parameters
Qs = range(0, 2)
qs = range(0, 3)
Ps = range(0, 3)
ps = range(0, 3)
D=1
d=1
parameters = product(ps, qs, Ps, Qs)
parameters_list = list(parameters)
len(parameters_list)

# Model Selection
results = []
best_aic = float("inf")
warnings.filterwarnings('ignore')
for param in parameters_list:
    try: model=sm.tsa.statespace.SARIMAX(df_month.Weighted_Price_box,
order=(param[0], d, param[1]),
seasonal_order=(param[2], D,
param[3], 12)).fit(dispatch=-1)
    except ValueError:
        print('wrong parameters:', param)
        continue
    aic = model.aic
    if aic < best_aic:
        best_model = model
        best_aic = aic
        best_param = param
    results.append([param, model.aic])

# Best Models
result_table = pd.DataFrame(results)
result_table.columns = ['parameters', 'aic']
print(result_table.sort_values(by = 'aic', ascending=True).head())
print(best_model.summary())

# STL-decomposition
plt.figure(figsize=(15,7))
plt.subplot(211)
best_model.resid[13:].plot()
plt.ylabel(u'Residuals')
ax = plt.subplot(212)

```

```

sm.graphics.tsa.plot_acf(best_model.resid[13:].values.squeeze(),
lags=48, ax=ax)

print("Dickey-Fuller test:: p=%f" %
sm.tsa.stattools.adfuller(best_model.resid[13:])[1])

plt.tight_layout()
plt.show()

# Inverse Box-Cox Transformation Function
def invboxcox(y, lambda):
    if lambda == 0:
        return(np.exp(y))
    else:
        return(np.exp(np.log(lambda*y+1)/lambda))

# Predição para períodos passados
df_month2 = df_month[['Close']]
date_list = [datetime(2017, 6, 30), datetime(2017, 7, 31),
datetime(2017, 8, 31), datetime(2017, 9,
30),
            datetime(2017, 10, 31), datetime(2017, 11, 30),
datetime(2017, 12, 31), datetime(2018, 1,
31),
            datetime(2018, 1, 28)]
future = pd.DataFrame(index=date_list, columns=df_month.columns)
df_month2 = pd.concat([df_month2, future])
df_month2['forecast'] = invboxcox(best_model.predict(start=0, end=75),
lambda)
plt.figure(figsize=(15,7))
df_month2.Close.plot()
df_month2.forecast.plot(color='r', ls='--', label='Predicted Close')
plt.legend()
plt.title('Bitcoin exchanges, by months')
plt.ylabel('mean USD')
plt.show()

# Predição para o presente momento
df_month2 = df_month[['Close']]
date_list = pd.date_range(start=df_month2.index[-1], periods=24,
freq='M')[1:]
future = pd.DataFrame(index=date_list, columns=df_month.columns)
df_month2 = pd.concat([df_month2, future])
df_month2['forecast'] = invboxcox(best_model.predict(start=0, end=99),
lambda)
plt.figure(figsize=(15,7))
df_month2.Close.plot()
df_month2.forecast.plot(color='r', ls='--', label='Predicted Close')
plt.legend()
plt.title('Bitcoin exchanges, by months')
plt.ylabel('mean USD')

```

```
plt.show()

# Predição para períodos futuros

df_month2 = df_month[['Close']]
date_list = pd.date_range(start='2017-06-30', end='2024-12-31',
freq='M')
future = pd.DataFrame(index=date_list, columns=df_month.columns)
df_month2 = pd.concat([df_month2, future]) df_month2['forecast']
= invboxcox(best_model.predict(start=0, end=len(df_month2)),
lambda)
plt.figure(figsize=(15,7))
df_month2.Close.plot()
df_month2.forecast.plot(color='r', ls='--', label='Predicted Close')
plt.legend()
plt.title('Bitcoin exchanges, by months')
plt.ylabel('mean USD')
plt.show()
```