



**A Probabilidade de Negociações Informadas vs *Disclosures* de Negociações dos *Insiders*:
Análise do caso da JBSS3**

GLAUCO GRACO NÓBREGA PORDEUS

Universidade Federal da Paraíba - UFPB

LUIZ FELIPE DE ARAUJO PONTES GIRÃO

Universidade Federal da Paraíba - UFPB

FILIPE COELHO DE LIMA DUARTE

Universidade Federal da Paraíba - UFPB

Resumo

Este trabalho tem como seu objetivo principal identificar operações de compra ou venda no ativo JBSS3 e introduzir uma metodologia de análise de operações com base na *proxy* de informação assimétrica captada pela PIN, adaptada por Lin e Ke (2011) e a ferramenta de *Machine Learning* desenvolvida por Chen e Guestrin (2016), XGboost, em seguida são comparadas tais operações com um modelo de negociação tendo como base as negociações feita pelos *insiders* e divulgadas ao mercado por meio do relatório referente ao Art. 11 da ICVM 358. Para realizar tal comparação foram propostas três hipóteses: H1, É possível obter retorno acima do mercado por meio de operações efetuadas com base nas informações obtidas no relatório referente ao Art. 11 da ICVM 358; H2, É possível obter retorno acima do mercado por meio de operações efetuadas com base na PIN e seus parâmetros; H3, É possível obter melhores rendimentos quando operando por meio da probabilidade de negociação informada do que por meio das informações presentes no relatório da ICVM 358. Os resultados encontrados para o ativo JBSS3, para o período de janeiro de 2016 até dezembro de 2018, apontam que a H1 foi rejeitada, a H2 foi parcialmente rejeitada e a H3, que buscou verificar e comparar as operações por meio do modelo da PIN com as informações da ICVM 358, não foi rejeitada. Assim, este estudo buscou aplicar conhecimentos computacionais ao mercado financeiro, como meio de promoção e aprimoramento da eficiência de mercado, neste caso, em específico ao ativo JBSS3.

Palavras chave: Informação assimétrica, *Insider trading*, Aprendizado de máquina, Retorno anormal.



1. INTRODUÇÃO

O objetivo principal do mercado financeiro é conectar agentes poupadores aos tomadores, assim promovendo uma alocação eficiente dos recursos. Um fator primordial para que este feito seja efetivado é a homogeneidade das informações presentes no meio, isso possibilita que a partir de variáveis comuns, os agentes tomem determinadas posições estratégicas no mercado.

A homogeneidade informacional é um dos preceitos para a eficiência de mercado, para o entendimento deste conceito, Fama (1970) propõe a Hipótese do Mercado Eficiente – HME onde iniciou seu pensamento de que o mercado pode ser eficiente em três formas, que diferem entre si pelo tempo em que preços absorvem as informações relevantes aos ativos, o autor classifica estas as formas da seguinte forma: forte, semiforte e fraca.

Ainda sobre o mercado financeiro e sua eficiência, Healy e Palepu (2001) afirmam que podem ocorrer falhas nos momentos de os negócios serem efetivados pelos agentes, poupadores e os tomadores. Estas falhas têm sua fundamentação em dois pontos, o primeiro, é no sentido da seleção adversa, onde os agentes compradores têm melhor conhecimento sobre as informações dos ativos focado, se comparados aos agentes vendedores; o segundo assinala o risco moral que os vendedores correm, pois existe a possibilidade de utilizar os recursos obtidos por meio da venda de seus ativos em benefício próprio, o que corrobora a Teoria da Agência (JENSEN; MECKLING, 1976).

Com base nisto, os autores Healy e Palepu (2001) ainda colocam a *disclosures* corporativas como um meio importante de disseminação de informações ao mercado sendo constituídas de previsões, relatórios da administração, demonstrações contábeis e afins. No Brasil o braço que atua neste sentido, com o propósito de regular, proteger a eficiência, a evolução e a seriedade do mercado de capitais brasileiro, é a Comissão de Valores Mobiliários (CVM), esta foi fundada em dezembro de 1976 a por meio da Lei 6.385/76, sendo caracterizada uma autarquia ligada ao Ministério da Fazenda do Brasil, dotada de autonomia financeira e orçamentária.

Um exemplo de falha da eficiência ocorreu no dia 17 de maio de 2017, em que após o fechamento do pregão deste dia veio à tona, por meio do portal da BBC, informações relevantes referentes à empresa brasileira JBS S.A., onde foi informado que os controladores estariam envolvidos em casos de corrupção em conjunto com agentes da União, informação que foi corroborada pela delação premiada divulgada pelo Supremo Tribunal Federal, dois dias depois, em 19 de maio de 2017 segundo o site BBC. Além disso, de acordo com o relatório referente ao art. 11 da Instrução CVM 358 (ICVM 358), estes controladores venderam parte de suas ações dias antes ao vazamento destas informações, o que pode ser um indício de uso de informações privadas para benefício próprio, por parte destes agentes. A tabela a seguir demonstra, de forma cronológica os principais fatos relacionados a JBS no ano de 2017.

A JBS S.A. é uma empresa brasileira do ramo de proteína animal, tem o capital aberto sendo listada na B3 por meio do *ticker* JBSS3 e é majoritariamente dividida entre a União e o grupo J&F Participações, este com 40% de suas ações, propriedade dos irmãos Wesley e Joesley Batista, já a União entra por meio do Banco Nacional de Desenvolvimento Econômico e Social – BNDES detendo 21,32% das ações, segundo suas informações presentes na B3.

Com melhor condição informacional, o mercado como um todo tem a capacidade de se ajustar de forma mais eficiente, visto que o *delay* de informação exigida pela CVM pode chegar a até 40 dias, pois os controladores têm obrigatoriedade de apresentar este relatório ao mercado até o décimo dia do mês subsequente ao referente (ICVM 358). Como exemplo, uma operação efetuada no dia 01/01 poderá chegar ao mercado apenas no dia 10/02. O uso da tecnologia de informação pode ser uma saída para estas anomalias.

No tocante ao desenvolvimento tecnológico, Tarouco e Graeml (2011) identificaram uma crescente preocupação das empresas no desenvolvimento de melhores práticas de governança de TI, colocando como um dos fatores determinantes desta busca, a exigência de transparência por parte do mercado.

O desenvolvimento tecnológico aplicado no mercado financeiro já é presente. O uso de *High Frequency Trading* - HFT em Wall Street, no ano de 2009, representava em torno de 60% do

volume negociado. Da parte dos operadores, o Medallion Fund obteve um rendimento médio de 35%, entre 2000 e 2010 (ALDRIDGE, 2010).

Com base no que foi exposto até o momento, e com a ciência de que a expansão tecnológica é necessária e contribui para a qualidade das informações no mercado, o presente estudo busca em seu objetivo geral, identificar se a Probabilidade de Negociação Informada – PIN tem a capacidade de antecipar as informações apresentadas pelos agentes da JBS no relatório da do Art. 11 da Instrução CVM 358 (ICVM 358).

Além disso, este estudo propõe (i) verificar se é possível obter retorno acima do mercado por meio das informações obtidas no relatório do Art. 11 da ICVM 358; (ii) identificar se é possível obter retorno acima do mercado por meio dos parâmetros obtidos no modelo da PIN adaptado por Lin e Ke (2011); (iii) analisar se é possível obter melhor desempenho financeiro (retorno) operando por meio da probabilidade de negociação informadas quando comparado com as operações efetuadas com base na ICVM358.

Para isso foi aplicado um modelo de aprendizado de máquina com o objetivo de identificar pontos de compra e venda no ativo JBSS3, tendo como base a assimetria informacional obtida pela *proxy* PIN, estas operações foram comparadas com as operações efetuadas pelos agentes referidos a ICVM 358. Dessa forma, os resultados apontaram que as operações efetuadas com base no modelo de aprendizado de máquina obtiveram desempenho superior às operações efetuadas com base nas negociações dos *insiders*.

2. REVISÃO DA LITERATURA E DESENVOLVIMENTO DAS HIPÓTESES

A literatura é vasta no que diz respeito à eficiência de mercado. Em Callado (2009) são encontradas diversas referências no que diz respeito ao início da construção do pensamento de eficiência de mercado, onde são citados os trabalhos de Campbell, Lo e MacKinlay (1997); Leroy (1989); e Ceretta (2001). Sendo exaltada a importância do estudo de Samuelson (1965), no qual foi introduzida a ideia de que os preços dos ativos têm movimentos aleatórios.

Na formulação da HME, Fama (1970) assume que a eficiência em sua forma forte, todas as informações, mesmo aquelas que ainda não são de domínio público, já se encontram precificadas; em sua forma semiforte, as informações são precificadas assim que são divulgadas ao público; por fim, em sua forma fraca, o autor coloca que as informações passadas ainda se mantêm absorvidas nos preços presentes e futuros.

Segundo Fama (1970), na HME, a função do mercado de capitais é disponibilizar um meio para que os recursos fluam entre os agentes superavitários e os agentes deficitários da economia. Seu funcionamento é baseado em três premissas chaves: (1) todos os agentes interessados em investir aceitam que os preços dos ativos sejam influenciados pelas *disclosures*; (2) as negociações são feitas sem custo de transação; e (3) todas as informações disponíveis são gratuitas para todos agentes do mercado.

Porém é trivial o entendimento de que por ser um modelo teórico, este representa uma aproximação da realidade, então a prática de suas três condições em conjunto não é efetivamente contínua. Neste sentido, Medeiros e Matsumoto (2002) constataram o uso de informações privilegiadas no momento da emissão de novas ações ao mercado, o que para agentes informados existe uma abertura à possibilidade de obtenção de retornos anormais.

No mercado de capitais, existem diversas formas de se aproximar aos indícios do uso de informações privilegiadas, algumas destas formas é a volatilidade dos retornos dos títulos, volume de negociações, comunicados ao mercado referentes à companhia, lucros extraordinários (DIERKENS, 1991); anúncio de emissão de ações e retornos anormais ao mercado (GARCIA, 2002). Outras formas de se identificar indícios de uso de informações não públicas, é por meio de modelos de microestrutura de mercado, como já citado neste trabalho.

A PIN é um destes modelos de microestrutura de mercado, foi desenvolvida por Easley et al. (2002) e reflete a assimetria a partir dos dados *intraday* das cotações do ativo analisado. Sua mensuração é captada por meio do desequilíbrio entre as operações de compra e de venda dos agentes presentes no mercado. Este é um modelo que dispõe aos analistas uma medida sólida do

grau de assimetria informacional para o período analisado (ABAD; RUBIA, 2005). No Brasil, Bosque (2016) desenvolveu uma abordagem bayesiana da PIN, com base no modelo aprimorado por Lin e Ke (2011), que, segundo Bosque (2016), esta abordagem possibilitaria introduzir opiniões dos analistas sobre os parâmetros da PIN.

Ainda no mercado brasileiro, Martins, Paulo e Girão (2014) ao analisarem o *value relevance* no modelo de Ohlson (1995) da assimetria de informação por meio da PIN de Easley et al. (2002) em 198 ações da Bolsa de Valores, Mercadorias e Futuros de São Paulo (atual B3), durante o ano de 2011, encontraram que a PIN agrega informação na avaliação de empresas, mais expressivamente em companhias que não se encontram em segmentos de governança corporativa reconhecidos. Os mesmos autores ao investigarem o possível caso de *insider trading* na OGX, encontraram que o preço da ação da companhia, entre os anos de 2008 e 2014, foi afetado positivamente pelas notícias no site da empresa e comunicados ao mercado e já quando relacionado aos fatos relevantes e formulários da ICVM 358, foi encontrada uma influência negativa. Com relação à PIN os autores ainda constataram que o mercado precificou o modelo de microestrutura até meados de 2011 o que, para este período, entrou em consonância ao proposto por Demsetz (1986), em relação à possibilidade de os portadores de informações privilegiadas obterem retornos anormais.

Desta maneira, levando em conta o potencial e a relevância do modelo da PIN proposta por Lin e Ke (2011) – PIN LK, seu desenvolvimento teórico e operacional no Brasil, e sua relevância em estudos internacionais onde foi adotado como *benchmark* nos estudos mais recentes relacionado a probabilidade de negociação com informação privada (LIN; KE, 2011; YAN; ZHANG, 2012; YAN; ZHANG, 2014; GAN; WEI; JOHNSTONE, 2015; ERSAN; ALICI, 2016; BOSQUE, 2016) o que fundamenta a utilização da PIN LK (2011) como *proxy* para assimetria informacional neste trabalho.

A CVM, no intuito de seguir seu objetivo regulador e garantidor da simetria informacional, em janeiro de 2002 instituiu a instrução 358 em que “disciplina a divulgação de informações na negociação de valores mobiliários e na aquisição de lote significativo de ações de emissão de companhia aberta”, mais especificamente em seu Art. 11 coloca que:

“Os diretores, os membros do conselho de administração, do conselho fiscal e de quaisquer órgãos com funções técnicas ou consultivas criados por disposição estatutária ficam obrigados a informar à companhia a titularidade e as negociações realizadas com valores mobiliários emitidos pela própria companhia, por suas controladoras ou controladas, nestes dois últimos casos, desde que se trate de companhias abertas”.

Ainda no diz respeito à possibilidade de obtenção de retorno anormal por parte dos detentores do capital, ao explorarem suas ações, Tonindael e Decourt (2012) encontraram, em empresas brasileiras, que estes agentes conseguem obter desempenho anormal tanto nas operações de compra, quanto nas operações de venda. Os autores constataram que para um mês após as compras dos *insiders*, houve retorno médio de 5,72%, e chegando até 9,87% seis meses após a operação. No que diz respeito às vendas dos *insiders*, houve retorno médio de -3,73%, chegando até -10,66% nos próximos seis meses. Quando relacionado PIN ao retorno das ações, Martins, Paulo e Albuquerque (2013) verificaram que em 198 empresas listadas na BMF&Bovespa (atual B3), o aumento em 10% da assimetria de informação medida pela PIN ocasiona um aumento no retorno das ações em 8%.

Este estudo está baseado no uso do algoritmo *eXtreme Gradient Boosting* (XGBoost) como ferramenta de classificação das operações de compra e venda. Este algoritmo foi desenvolvido por Chen & Guestrin (2016) e é, segundo os autores, amplamente usado por *data scientists* para obtenção de seus resultados em competições de aprendizado de máquina (*machine learning*) e foi vencedor do prêmio John Chambers Award (2016) em estatística computacional.

O XGBoost é um método de classificação com base no conceito de árvores de decisão, porém com diferenças significantes. Este modelo basicamente treina as novas árvores de decisão com base em suas estruturas fracas anteriores. Inicialmente são modeladas regressões para explicar os dados, de forma randômica, obtendo seus erros, e em seguida, com base nestes erros, são

desenvolvidas novas regressões onde é extraído um novo modelo com melhor desempenho (*i.e.*, melhor capacidade preditiva). Desta forma, este processo segue os seguintes passos: (1) aprendizagem do modelo preditivo; (2) computação dos resíduos; e (3) aprendizagem para predição dos resíduos (DEY; *et al.*, 2016).

Este modelo foi usado em aplicações no mercado financeiro para previsão das tendências das ações, no qual teve um acerto de 87% para os períodos de 60 e 90 dias (DEY; *et al.*, 2016). Yu (2017) utilizou esta ferramenta para prever o *Chicago Board Options Exchange (CBOE) Volatility Index – VIX*, do primeiro até o sexto pregão subsequentes da análise, obtendo uma taxa de acerto entre 55% e 65%, dependendo da janela analisada, sendo a melhor acurácia obtida aplicando o modelo para 5 e 6 dias.

Sob outra perspectiva, Carmona *et al.* (2018) utilizou o algoritmo XGBoost com o objetivo de prever a falência de bancos no mercado dos Estados Unidos entre os anos de 2001 e 2015, com uma série de 30 indicadores financeiros de 156 bancos comerciais, obtendo um acurácia de 94,74%, ou seja, o melhor desempenho quando comparado aos demais modelos testados em seu trabalho, a Regressão Logística e *Random Forest* que obtiveram acertos de 84,21% e 92,11% respectivamente.

Neste sentido, com o incentivo de ampliar os estudos na área e testar a ferramenta XGBoost no mercado financeiro, são estabelecidas as hipóteses (H1, H2 e H3) deste estudo:

H1: É possível obter retorno acima do mercado por meio de operações efetuadas com base nas informações obtidas no relatório referente ao Art. 11 da ICVM 358;

H2: É possível obter retorno acima do mercado por meio de operações efetuadas com base na PIN e seus parâmetros;

H3: É possível obter melhores rendimentos quando operando por meio da probabilidade de negociação informada do que por meio das informações presentes no relatório da ICVM 358.

3. MÉTODO

Para identificar se houve retorno acima de mercado nas operações do papel JBSS3 durante os anos 2016 e 2017, foram simuladas entradas no papel com ajustes mensais, de acordo com a divulgação dos relatórios da ICVM 358. Este estudo se limita a analisar apenas dois anos pelo fato de que os dados necessários para o cálculo da PIN, disponibilizados pela B3, estão restritos aos dois anos antecedentes a data da consulta.

Os dados, referente às negociações, utilizados para o cálculo do modelo da PIN e dos retornos foram extraídos por meio do programa *R*, onde principalmente foi utilizado o pacote *GetHFDData* (PERLIN; RAMOS, 2016) para *download* e tabulação destes dados de negociações diárias do período analisado. Para o cálculo do modelo da PIN foi utilizado o *script* disponibilizado por Bosque (2016) em seu apêndice C – PROGRAMAÇÕES UTILIZADAS.

Já para a classificação das operações de compra e venda, identificadas pela PIN e seus parâmetros, foi aplicado o algoritmo XGBoost, e para tanto, foi utilizado o pacote do *R*, denominado de *xgboost*, que é atualizado e mantido pelos próprios autores do algoritmo.

3.1 Estimação da Probabilidade de Negociação Informada

A mensuração da proxy de assimetria de informação nas cotações da JBSS3 foi captada por meio da probabilidade de negociação com informações privilegiadas (PIN) de Easley, Hvidkjaer e O'Hara (2002; 2010) adaptada por Lin e Ke (2011). Esta é obtida a partir do número e sentido dos negócios das ações (ordens de compra e venda) em determinado período de tempo. O modelo da PIN desenvolvida por EHO identifica uma quantidade recorrente de envio de ordens de compra (ϵ_b) e venda (ϵ_s) das ações, e toma como base para sinalizar as negociações dos negociantes desinformados, então um nível anormal de compras e vendas é utilizado para classificar negociações dos agentes informados (μ). A probabilidade de uma negociação de agentes informados (α) é obtida por meio do número de dias em que o nível de negociação foi anormal. A partir disso, a PIN é estimada por meio da união desses fatores simultaneamente, como demonstra a Equação (1).

$$PIN = \frac{\alpha\mu}{\alpha\mu + \epsilon_b + \epsilon_s} \quad 1)$$

Em que:

α é a probabilidade de ocorrer um evento informacional;

μ é taxa de negociações de agentes informados;

ε_b é a taxa de compra de agentes desinformados; e

ε_s é a taxa de venda de agentes desinformados.

Bosque (2016) coloca que os autores Lin e Ke (2011) propuseram um arranjo para o melhoramento do modelo citado anteriormente, essa melhoria foi instigada pela presença de imprecisões no processo computacional da PIN de EHO (2002; 2010). Além disso, outros trabalhos corroboram e abordaram o mesmo ponto no que diz respeito à falha do mesmo modelo, este trabalho são: Yan e Zhang (2012), Gan Wei e Johnston (2015) e Ersan e Alice (2016). Neste sentido foi escolhido para este estudo o modelo adaptado da PIN por LK (2011), pelo fato dele ser atualmente, um benchmark para novos estudos (BOSQUE, 2016).

Assim, segue o método abordado para o modelo LK onde, pós-classificação e apuração das transações de compra e venda, são estimados os parâmetros do modelo apresentados na Equação (1), por meio de um modelo de negociação sequencial. Tal estimativa será desenvolvida pela maximização de uma função máxima verossimilhança (Equação 2)

$$L(\theta|B, S) = \sum_{i=1}^n \left\{ \log \left[\frac{\alpha \delta \exp(e_{1i} - e_{\max i}) - \alpha(1 - \delta) \exp(e_{2i} - e_{\max i})}{(1 - \delta) \exp(e_{3i} - e_{\max i})} \right] + B_i \log(\varepsilon_b + \mu) + S_i \log(\varepsilon_s + \mu) \right. \\ \left. - (\varepsilon_b - \varepsilon_s) + e_{\max i} - \log(B_i! S_i!) \right\},$$

Sendo:

$$e_{1i} = -\mu - B_i \log \left(1 + \frac{\mu}{\varepsilon_b} \right) \quad (2)$$

$$e_{2i} = -\mu - S_i \log \left(1 + \frac{\mu}{\varepsilon_s} \right)$$

$$e_{3i} = -\mu - B_i \log \left(1 + \frac{\mu}{\varepsilon_b} \right) - \mu - S_i \log \left(1 + \frac{\mu}{\varepsilon_s} \right)$$

$$e_{\max i} = \max(e_{1i}, e_{2i}, e_{3i}), \quad i = 1, 2, \dots, n$$

Em que:

B e S são os volumes de compras e vendas;

B_i e S_i são as negociações no dia i;

θ é o vetor de parâmetros (α , μ , δ , ε_b e ε_s);

$(1-\delta)$ é a probabilidade de ser uma “boa notícia”; e

δ é a probabilidade de ser uma “má notícia”.

3.2 Critérios das Operações

As operações foram segmentadas de duas formas, a primeira por meio das informações coletadas na Art. 11 da ICVM 358, a segunda foi por meio da PIN e seus parâmetros (α , δ , μ , ε_b e ε_s). Tendo como base a fundamentação teórica de Grossman e Stiglitz (1980), que apontam a ineficiência de mercado na precificação dos ativos, porém as ações dos agentes informados podem contribuir para disseminação das informações por meio dos preços dos ativos, no Brasil, o trabalho de Martins, Paulo e Girão (2015) ressalta a importância da transparência das informações e reafirma a proposta de Grossman e Stiglitz (1980) no tocante ao acompanhamento das operações dos insiders.

Neste sentido, a identificação das operações por meio da ICVM foi estabelecida se baseando nos critérios adotados por Tonindael e Decourt (2012), que selecionaram as operações com volumes financeiros mais expressivos e relevantes. Seguindo este pensamento, foram estabelecidas duas formas de operar: (1) quando os membros referidos na ICVM 358 negociarem 1% ou mais de suas ações, seguindo o sentido da operação (venda ou compra); (2) quando os

membros referidos na ICVM 358 negociarem 0,5% ou mais de suas ações, seguindo o sentido da operação (venda ou compra).

Para as operações com base na PIN, será utilizado o algoritmo XGBoost, como já afirmado neste estudo. O input deste modelo preditivo será os *outputs* do modelo LK, sendo eles a própria PIN, α , δ , Negociações Informadas e Negociações Desinformadas (ϵ_b e ϵ_s); tendo como variável depende os retornos observados acima de mercado, sendo ajustados para saídas binária (0 ou 1), computando 1 quando os retornos forem maiores do que o mercado, e 0 para retornos abaixo do mercado, o que leva, portanto, a formação de um modelo de classificação. Para *proxy* de mercado foi estabelecido o Ibovespa, o qual segundo a Brasil, Bolsa, Balcão (B3), é o índice que concentra as maiores e mais líquidas empresas do mercado brasileiro.

Os retornos ajustados ao mercado são expressos da seguinte maneira, como demonstra a Equação (3):

$$Ra = \left(\frac{Pa_{tf}}{Pa_{t0}} - 1 \right) - \left(\frac{Pm_{tf}}{Pm_{t0}} - 1 \right) \quad (3)$$

Em que:

Ra é o retorno do ativo ajustado ao mercado;

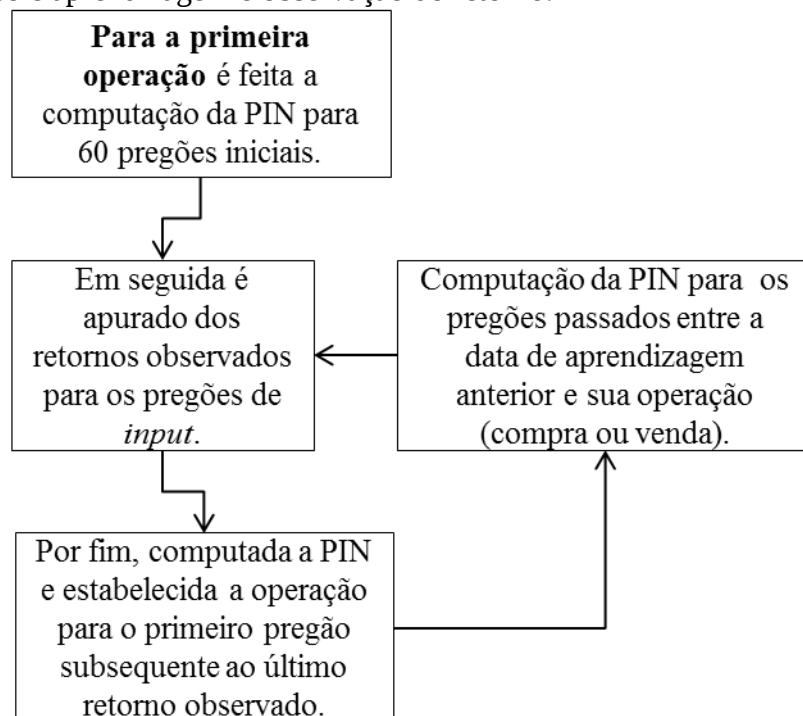
Pa_{tf} é preço do ativo na data futura;

Pa_{t0} é o preço do ativo na data de estimação da PIN;

Pm_{tf} é índice proxy do mercado na data futura; e

Pm_{t0} é índice proxy do mercado na data estimação da PIN.

Após a computação dos valores da PIN, seus parâmetros e do retorno ajustado ao mercado para os dias da amostra, que consiste nos pregões entre os dias 13/01/2016 e 08/12/2017 (totalizando 475 pregões), é realizada a primeira aprendizagem (estimação) do modelo para classificação da primeira operação (compra ou venda), aplicando o XGboost tendo como entrada para aprendizagem os dados dos 60 primeiros pregões da amostra, a Figura (1) seguinte demonstra o fluxo do processo e aprendizagem e observação do retorno.



Fonte: Elaborado pelo autor

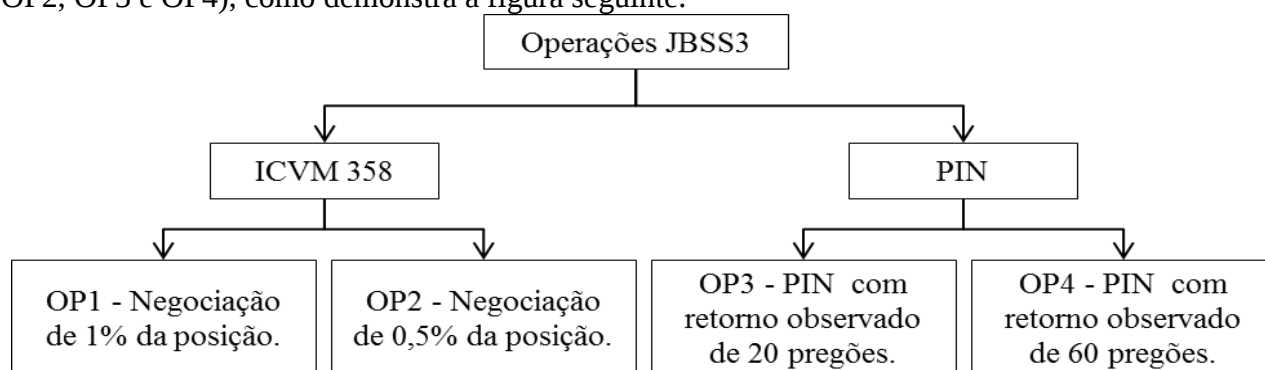
Figura 1: Rotina de aprendizagem e predição do XGBoost

Neste estudo, se optou por testar duas janelas de entrada para a variável dependente (retorno ajustado ao mercado). Portanto, foi desenvolvida a análise para dois modelos distintos: o

primeiro modelo é observado os retornos ajustados ao mercado no 20º pregão subsequente à apuração dos parâmetros da PIN; enquanto que o segundo modelo tem sua temporalidade ampliada para o 60º pregão posterior, com objetivo de captar o *delay* em relação a reação do mercado frente ao cenário nada data de apuração da PIN, visto que Tonidael e Decourt (2012) encontram que, testando os mesmos espaços temporais, quanto maior o período de observação, maior é o retorno anormal.

Com relação aos parâmetros adotados no XGBoost, o modelo é utilizado por meio do comando *xgboost*, do pacote com o mesmo nome, tendo a saída por meio de regressão logística binária, com valores probabilísticos entre 0 e 1. Dessa forma, para os valores acima de 0,5 foi adotado como operações de compra, e para os menores que 0,5 como operações de venda. Foi aplicado o comando *early stopping* para parar as simulações automaticamente quando encontrado um valor mínimo para o erro.

Como forma de ilustração, a Figura 2 representa o fluxograma de extração das operações por meio dos quatro modelos de decisão aplicados neste estudo. Na descrição dos resultados deste trabalho, serão apresentados os frutos dos modelos, denominados de operações 1, 2, 3 e 4 (OP1, OP2, OP3 e OP4), como demonstra a figura seguinte.



Fonte: Elaborado pelo autor

Figura 2 – Fluxo do Processo Decisório das Operações

A data das operações é referente à data de divulgação dos relatórios da ICVM 358, portanto, a cada divulgação é analisado se será efetuada compra, venda ou não operar, para ocasiões que não houverem informações relevantes para a fundamentação de uma posição.

4. RESULTADOS

O tratamento e computação dos dados foi feita com o auxílio do programa R. A Tabela (1) seguinte demonstra a distribuição da amostra em seus valores mínimos, máximos, médios, e em seu primeiro, segundo e terceiro quartil.

Tabela 1 – Estatística descritiva da amostra para 2016 e 2017

	Re 20p ¹	Re 60p ²	ICVM ³	PIN	α	δ	Eb	Es	Mu
Mínimo	-48,49%	-58,52%	0%	8,54%	13,33%	0%	6694	6188	4276
1º quartil	-12,28%	-18,84%	0%	12,26%	26,53%	26,21%	7645	7733	7481
Mediana	-2,95%	-7,50%	0%	15,10%	33,33%	40,01%	8524	8521	9843
Média	-5,07%	-8,99%	0,25%	15,13%	32,96%	44,90%	9243	8715	10471
3º quartil	5,97%	7,58%	0,03%	17,66%	38,33%	57,63%	9634	9479	12238
Máximo	27,06%	26,63%	2,70%	22,54%	58,81%	100%	15876	11631	21662

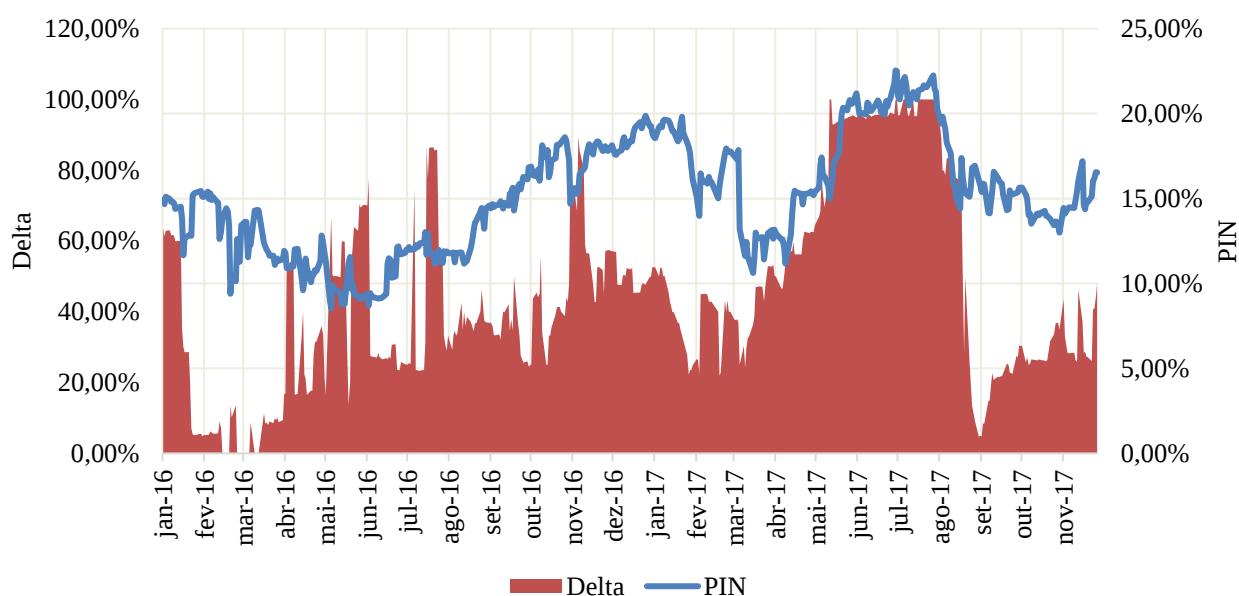
Fonte: Elaborado pelo autor. ¹ refere-se aos retornos no 20º pregão subsequente ao da apuração da PIN; ² refere-se aos retornos no 60º pregão subsequente ao da apuração da PIN; ³ refere-se ao volume (%) de negociação da posição dos agentes.

A partir da Tabela 1, pode-se observar que, em média, os retornos ajustados ao mercado são negativos, em -5,07% e -8,99%, para os intervalos de 20º e 60º pregão posterior, respectivamente, e que, com exceção do terceiro quartil, o retorno ajustado ao mercado em todos os intervalos da amostra para o 60º pregão posterior são inferiores ao retorno ajustado da amostra do 20º pregão, o que indica uma tendência de baixa para o ativo da JBSS3, em um horizonte de tempo maior. Com

relação às informações obtidas por meio da ICVM 358, observa-se uma baixa movimentação dos *insiders* visto que 75% dos meses os *insiders* negociaram menos que 0,03% de sua posição total (este ponto será abordado posteriormente na apresentação da Tabela 4 deste trabalho). De acordo com o modelo LK, a probabilidade diária de haver negociação uma negociação informada, para os anos de 2016 e 2017 no ativo JBSS3, foi em média 15,13%, valor consistente com o observado com trabalhos anteriores sobre o assunto, como observado por Martins, Paulo e Girão (2013). Com relação ao Delta, em média, a chegada de notícias é positiva, o que é representado por um delta abaixo de 50%, informação corroborada pela sua média (44,90%) e pelo número de negociações efetuadas pelos dos agentes compradores, em relação aos agentes vendedores, tendo em média um número de negociações 8,54% superior, quando observado os intervalos da amostra apresentada na Tabela 1,

Com intuito de promover uma melhor visualização da relação entre a probabilidade de negociação informada e a probabilidade desta informação ser positiva ou negativa está descrita abaixo, no Gráfico 1:

Gráfico 1: Relação *intraday* entre PIN e o Delta entre janeiro de 2016 e dezembro de 2017



Fonte: Elaborado pelo autor.

A média da PIN para o período analisado foi 15,13%, como visto na Tabela 1, e com base nisto, é possível notar que nos meses anteriores a delação, 01 de novembro de 2016 e 17 maio de 2017 (data do vazamento da delação), o ativo JBSS3 teve uma PIN média de 17,79%, valor 17,55% superior à média total do período, neste mesmo período em que a PIN estava acima da média, o sentido do delta se manteve praticamente neutro, com uma média de 46,40%, o que é 3,35% superior a media total da amostra, e simboliza incerteza com relação ao sentido da negociação informada.

Outro período de destaque está entre o começo de julho e o final de agosto de 2017, quando a PIN média foi de 20% e o Delta médio de 90,49%, ficando num patamar superior à média total da amostra em 32,20% e 101,54%, respectivamente, fatores que demonstram uma alta probabilidade de negociação com informação negativa, privada ao mercado, fato que pode ser relacionado ao pronunciamento feito pelo então procurador-geral da República Rodrigo Janot, no dia 4 de setembro de 2017, direcionando a possibilidade cancelamento dos benefícios da delação premiada feita pelos controladores da companhia.

Já no que diz respeito a negociação dos *insiders*, representada pelas informações obtidas no relatório da ICVM 358, a atuação mais relevante destes agentes é efetivada durante os meses de maio, junho e julho de 2017, onde eles negociaram ao todo 85,6% das ações do total analisado, entre 2016 e 2017.

As tabelas 1 e 2, a seguir, apresentam os fatos ocorridos durante o período antes e depois da



delação premiada dos agentes da JBS, os meses em que houve movimentações nas posições dos agentes, a data de apresentação do relatório ao mercado, sentido da negociação, posição negociada em relação ao total e o quanto a negociação representa ao total do período analisado.

Tabela 1 - Cronologia dos fatos relacionados à delação da JBS de janeiro até dezembro de 2017

Fato	Data
Operação Cui Bonno	13 de janeiro de 2017
Marcelo Miller pede exoneração	23 de fevereiro de 2017
Temer recebe Joesley	07 de março de 2017
Operação carne fraca e diálogo com Ricardo Saud	17 de março de 2017
Joesley grava Aécio Neves	24 de março de 2017
Negociação com a Procuradoria-Geral	27 de março de 2017
Marcelo Miller deixa o Ministério público	05 de abril de 2017
Ações controlada e depoimentos	Em abril de 2017
Assinatura da delação	03 de maio de 2017
Homologação da delação	10 de maio de 2017
Vazamento da delação	17 de maio de 2017
Aécio Neves é denunciado	02 de junho de 2017
Michael Temer é denunciado	27 de junho de 2017
Câmara barra denúncia de Aécio Neves	02 de agosto de 2017
Acordo de delação pode ser anulado	04 de setembro de 2017

Fonte: adaptado de Nexo

Tabela 2 – Negociação dos *insiders* apresentadas por meio da ICVM 358, entre janeiro e dezembro de 2017

Data de entrega do relatório ao mercado	Compra	Venda	Quantidade de ações negociada	Proporção das negociações em relação ao período	Posição negociada (%)
10/02/2016	191mil	197,2mil	388,2mil	0,56%	0%
10/03/2016	1.169mil	1.227,7mil	2.396,7mil	3,43%	0%
08/04/2016	-	63,1mil	63,1mil	0,09%	0%
10/05/2016	-	10mil	10mil	0,01%	0%
09/01/2017	-	40mil	40mil	0,06%	0%
09/01/2017	-	10mil	10mil	0,01%	0%
08/03/2017	-	254,5mil	254,5mil	0,36%	0%
10/05/2017	-	31.777,5mil	31.777,5mil	45,46%	2,63%*
10/06/2017	-	18.635,7mil	18.635,7mil	26,66%	1,58%*
10/07/2017	-	9.413,3mil	9.413,3mil	13,47%	0,81%*
10/08/2017	-	385,5mil	385,5mil	0,55%	0,03%
08/09/2017	-	6.520mil	6.520mil	9,33%	0,57%*
10/11/2017	-	1mil	1mil	0%	0%

Fonte: Elaboração pelo autor. * Meses em que a negociação foi fora da normalidade, caracterizando operação de acordo com o modelo 1 e/ou 2 referentes a ICVM 358.

Desta maneira, conforme o que foi exposto até o momento, as operações com a ICVM 358 foram simuladas para de maio, junho, julho e setembro de 2017, seguindo o sentido das operações dos *insiders*. Para a PIN foram simuladas entradas em 11 meses com o modelo com retorno observado no 20º pregão subsequente e em 9 ocasiões para o modelo com retorno observado no 60º pregão subsequente, já para operações de venda foram 9 para o modelo com retorno observado no 20º pregão subsequente e 8 para o modelo com retorno observado no 60º pregão subsequente.

A tabela a seguir compara o resultado das simulações, incluindo o Ibovespa.

Tabela 3 – Comparativo das operações efetuadas com informações de *insiders* entre janeiro de 2016 e dezembro de 2017

	Ibovespa	ICVM 1%	ICVM2 0,5%	PIN 20p	PIN 60p
Meses Analisados	-	24	24	24	24
Sem Operações	-	22*	20*	4**	7**
Compras	-	0	0	11	9
Vendas	-	2	4	9	8
Taxa de Acerto***	-	100,0%	50,0%	60%	64,71%
Retorno Total	91,79%	49,49%	10,8%	78,13%	127,84%
Retorno Médio	3,82%	2,06%	0,45%	3,26%	5,33%
Desvio-padrão	6,15%	8,03%	9,56%	14,42%	12,92%

Fonte: Elaboração pelo autor. * não houve operação significativa dos *insiders*; ** refere-se aos meses em que foram computadas as entradas para o modelo XGBooster; *** refere-se a relação entre o que foi proposto pelos modelos sobre o observado no mercado.

As operações, quando efetuadas com base na ICVM 358, obtiveram retorno positivo, sendo que, não é possível assumir a significância destas operações, visto que dos 24 períodos observados foram identificadas apenas 2 operações referentes ao primeiro modelo de seleção, quando os agentes previstos na ICVM 358 negociarem mais que 1% de suas ações, que obtiveram 100% de acerto. Vale ressaltar que o período em que este modelo sinalizou para operar foi nos meses maio e abril, meses subsequentes ao vazamento da delação premiada dos seus controladores. Já para o segundo modelo houve mais pontos de operações, quatro ao todo, sendo que destas, 50% foram operações falhas, ou seja, com perdas.

No que diz respeito à rentabilidade destes modelos, o primeiro obteve um retorno total de 49,49%, já o segundo 10,8%, sendo que para o mesmo período o Ibovespa, índice adotado como representação do mercado, obteve uma rentabilidade de 60,62%, o que significa que ambos os modelos obtiveram um desempenho inferior ao mercado em -11,13% e -49,82, respectivamente. Desta maneira a H1, que buscou verificar se seria possível obter retorno acima do mercado por meio de operações efetuadas com base nas informações obtidas no relatório referente ao Art. 11 da ICVM 358, é rejeitada.

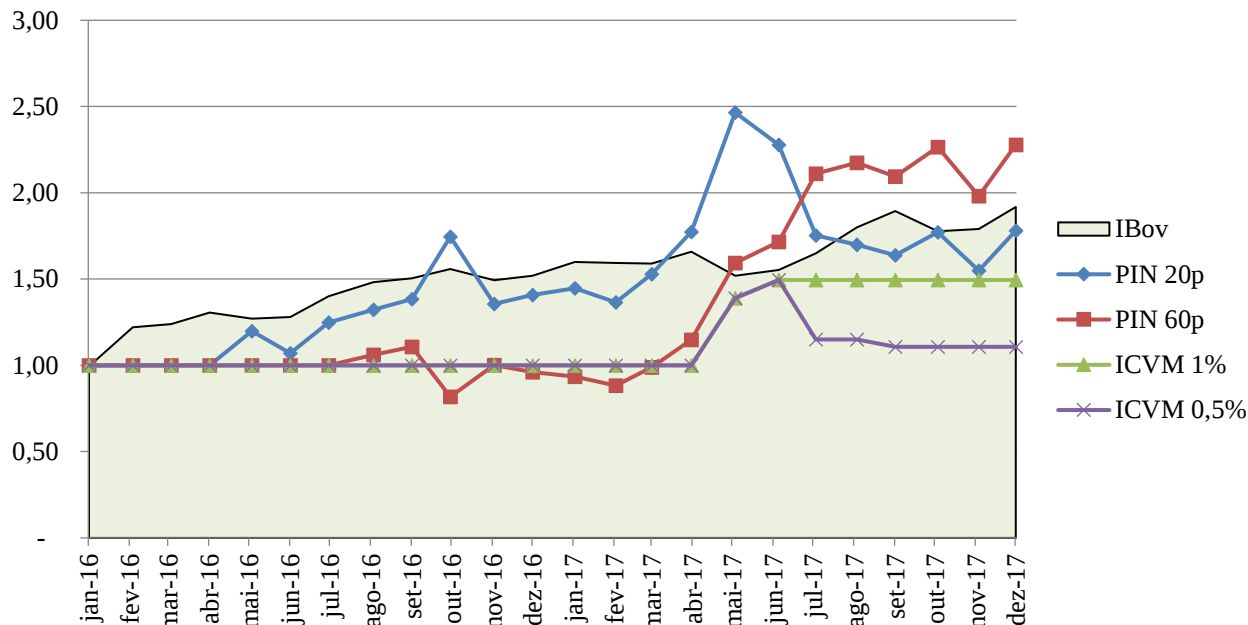
Já com relação às operações indicadas pelo XGboost, tendo como base a PIN, foram efetuadas 20 operações, para o modelo com retorno no vigésimo pregão subsequente, e 17 operações para o modelo com retorno no sexagésimo pregão subsequente, acertando, respectivamente, 12 das 20 operações (60%), 11 das 20 operações simuladas (64,71%).

Já no tocante à rentabilidade destes modelos, o modelo com análise do 20º pregão obteve um retorno total de 78,13%, já o segundo obteve um retorno observado de 127,84%. Assim, apenas um dos modelos com base na PIN obtiveram retorno acima do Ibovespa, o modelo com maior janela de retorno, portanto não foi possível confirmar H2, pois não foi possível obter retorno acima do mercado por meio de operações efetuadas com base na PIN e seus parâmetros com todos os dois modelos testados.

Já a H3 não é rejeitada, visto que ambos os modelos com base na PIN, nesta amostra estudada, foram possíveis obter melhores rendimentos quando operando por meio da probabilidade de negociação informada do que por meio das informações presentes no relatório da ICVM 358.

Por fim, com o objetivo de visualizar melhor o desenvolvimento dos modelos durante o período, está disposto o gráfico 2, onde se comparam todos os modelos abordados neste estudo e o desenvolvimento do Ibovespa, adotando como o início da série o valor de uma unidade monetária, para efeito de capitalização. É relevante ressaltar que ambos os modelos da PIN, a partir de fevereiro de 2017, meses antes do vazamento da delação, apresentaram alta expressiva.

Gráfico 2 - Capitalização dos modelos e comparativo com o Ibovespa, entre 2016 e 2017



Fonte: Elaborado pelo autor.

5. CONSIDERAÇÕES FINAIS

Este estudo teve com objetivo identificar operações com base em negociações com a probabilidade de negociações informadas, para tanto aplicou-se o modelo de aprendizado de máquina aos outputs da PIN, obtendo um retorno acima do observado das operações efetuadas pelos agentes internos à companhia analisada, JBS AS.

Os resultados encontrados para o período analisado constatarem que as operações efetuadas por meio do modelo XGBoost obtiveram um resultado superior às operações efetuadas pelos os *insiders*, tendo em vista que de acordo com a ICVM 358, estas operações são influenciadas pelo *delay* entre as reais operações dos agentes internos à companhia e a disseminação desta informação ao mercado, visto que os *insiders* têm até o décimo dia útil do mês subsequente para informar as operações efetuadas, o que pode comprometer o desempenho daqueles que operam seguindo suas operações (o que demonstra as simulações com base nesta informação).

Vale destacar que este estudo está limitado ao ativo da JBSS3 para o período de janeiro de 2016 até dezembro de 2017, visto que a disponibilidade dos dados de cotações no repositório *online* da B3 está em uma janela de dois anos, limitando o trabalho por acesso a este tipo de informação. Desta maneira, para melhor verificação da relevância do modelo adotado, sugere-se ampliar a amostra para os demais ativos da bolsa em estudos futuros.

Ademais, é importante ressaltar a relevância do avanço da tecnologia de inteligência de máquina e seu acesso mais intuitivo, a união deste nicho de conhecimento com demais áreas da ciência, inclusive com o mercado de capitais, é fundamental para aperfeiçoamento do processo de tomada de decisão e manuseio de novas fontes de informações.

REFERÊNCIAS

ABAD, D.; RUBIA, A. Modelos de estimación de la probabilidad de negociación informada: unacomparación metodológica en el mercado Español. **Revista de Economía Financeira**, n. 7, p. 1- 37, 2005.

BBC. **Em gravação feita por dono da JBS, 'Temer dá aval a compra de silêncio de Cunha', diz jornal; presidente nega**. Disponível em: <<http://www.bbc.com/portuguese/brasil-39956081>> Acesso em: 22 de fevereiro de 2018



BBC. **Guia da delação da JBS: entenda as acusações que abalaram o mundo político.**

Disponível em: <<http://www.bbc.com/portuguese/brasil-39983080>> Acesso em: 23 de fevereiro de 2018

CALLADO, A. A. C. Eficiência do mercado acionário brasileiro: retorno das ações negociadas na Bovespa, variáveis macroeconômicas, causalidade e fatores condicionantes. Tese (Doutorado). Programa de Pós-Graduação em Administração, Universidade Federal de Pernambuco, Recife, 2009.

CARMONA, P., CLIMENT, F., & MOMPALER, A. (2018). Predicting bank failure in the US banking sector: An extreme gradient boosting approach. **International Review of Economics & Finance**.

CHEN, T., & GUESTRIN, C. Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794). ACM, 2016.

ALDRIDGE, I. High-Frequency trading: a practical guide to algorithmic strategies and trading systems. New Jersey: **John Wiley & Sons**, 2010.

ANTUNES, T. F. A utilização de robôs de negociação em alta frequência na obtenção de ganhos financeiros e suas consequências para o mercado, 2015.

BOSQUE, L. M. Estimação da probabilidade de negociação privilegiada por meio de inferência bayesiana. 2016. 141 f. Dissertação (Mestrado em Administração: Finanças). Universidade de Brasília. Brasília. 2016.

ASLEY, D.; HVIDKJAER, S.; O'HARA, M. Is information risk a determinant of asset returns? **Journal of Finance**. v.57. n.5. p.2185-2221. 2002.

DEY, S., KUMAR, Y., SAHA, S., & BASAK, S. Forecasting to Classification: Predicting the direction of stock market price using Xtreme Gradient Boosting. Working paper. DOI: 10.13140/RG.2.2.15294.48968.

EASLEY, D.; O'HARA, M. Information and the cost of capital. **The Journal of finance**, v. 59, p. 1553-1583, 2004.

EASLEY, D; KIEFER, N. M.; O'HARA, M.; PAPERMAN, J. B. Liquidity, information and infrequently traded stocks. **The Journal of Finance**, v. 51, n. 4, p. 1405-1436, 1996.

FAMA, E. F. Efficient capital markets: a review of theory and empirical work. **The journal of finance**, v. 25, n. 2, p. 383-417, 1970.

GROSSMAN, S. J., & STIGLIZ, J. E. On the impossibility of informationally efficient markets. **The American Economic Review**, 70(3), 393-408, 1980.

GIRÃO, L. F. A. P.; MARTINS, O. S.; PAULO, E. Insider trading B-side: relevance, timeliness and position influence. **Revista Brasileira de Gestão de Negócios**, v. 17, n. 58, p. 1341-1356, 2015.

HEALY, P. M.; PALEPU, K. G. Information asymmetry, corporate disclosure, and the capital markets: a review of the empirical disclosure literature. **Journal of Accounting and Economics**, v. 31, p. 405-440, 2001.

JENSEN, M. C.; MECKLING, W. H. Theory of the firm managerial behavior, agency costs and ownership structure. **Journal of Financial Economics**, v. 3, n. 4, p. 305-360, 1976.

John Chambers Award. **Past winners**. Disponível em: <<http://stat-computing.org/awards/jmc/winners.html>> Acesso em: 23 de fevereiro de 2018.

Jornal Nexo. <<https://www.nexojournal.com.br/expresso/2017/09/05/A-cronologia-do-caso-JBS-press%C3%A3o-dela%C3%A7%C3%A3o-e-questionamentos>> Acesso em: 24 de fevereiro de 2018



MARTINS, O. S.; PAULO, E. A probabilidade de negociação com informação privilegiada no mercado acionário brasileiro. **Revista Brasileira de Finanças**, v. 11, n. 2, p. 249-280, jun. 2013.

MARTINS, O. S.; PAULO, E.; ALBUQUERQUE, P. H. Negociação com informação privilegiada e retorno das ações na BM&FBOVESPA. **RAE – Revista de Administração de Empresas**, São Paulo, v. 53, n. 4, p. 350-362, jul./ago. 2013.

PERLIN, M., RAMOS, H. GetHFData: A R Package for Downloading and Aggregating High Frequency Trading Data from Bovespa. **Brazilian Review of Finance**, V. 14, N, 2016.

TAROUCO, H. H.; GRAEML, A. R. Governança de tecnologia da informação: um panorama da adoção de modelos de melhores práticas por empresas brasileiras usuárias. **Revista de Administração**, v. 46, n. 1, p. 7-18, 2011.

TONINDAEL, M. C., & DECOURT, R. F. Insider Trading: um estudo sobre a rentabilidade das operações com ações da própria empresa. **Anais do Congresso Virtual Brasileiro - Administração**, 2011.

YU, M. Y. (2017). Predicting the Volatility Index Returns Using Machine Learning (Doctoral dissertation).