

Identificação em tempo real de Manipulação no Mercado Financeiro de Ações e Derivativos por Spoofing e Layering

Mateus Saldanha⁽¹⁾ Dr. Nizam Omar⁽²⁾

Universidade Presbiteriana
Mackenzie

rlhsaldanha@gmail.com⁽¹⁾ omar@mackenzie.br⁽²⁾

ABSTRACT

The manipulation in the stock market is foreseen and described by the main regulators of the financial market, to avoid that someone can use insider information, or even technological, financial or operational subterfuges, try or distort negotiation data to obtain undue advantage. This work aims to propose an autonomous and online form, using Machine Learning to identify disruptive movements in the negotiation from the use of spoofing and layering, helping in the quick identification and selection of purchase and sell orders that have generated or were generated by this movement.

Keywords

Manipulação, Spoofing, Layering, Machine Learning, kMeans, Apriori, Adaboost, Regulação, Real Time.

1. INTRODUÇÃO

O Mercado de Ações está em processo de desenvolvimento e da mesma forma a sua regulação. Importantes medidas reguladoras como a Lei Americana Dodd-Frank (2016) [13] tem contribuído na atualização da regularização e acompanhamento do mercado.

A Lei Americana, em seu estatuto, aponta quatro possíveis exemplos de situações classificadas como Spoofing, premissas semelhantes adotadas no mercado brasileiro, pela CVM na normativa publicada em 1979[10].

O termo “disruptive trading behavior”, é atualmente utilizado para descrever atividades de negociação que criam uma falsa imagem de liquidez do mercado, profundidade no book de negociação, ou ainda, uma condição artificial de movimento do preço, seja de aumento ou redução, de um determinado instrumento, que não esteja aderente ao instrumento negociado[12], conforme citado, no trabalho Computational intelligent hybrid model for detecting disruptive trading activity [1], e primeiramente descrito pela US Commodity Futures Trading Commission.

Em virtude da variedade de artigos na língua inglesa, este trabalho manterá os termos mais utilizados em inglês sem a tradução para o português, para aproximar os dados com base nas referências abordadas e dos termos utilizados pelo mercado.

2. REFERENCIAL TEÓRICO

A manipulação do mercado acionário, segue a definição dos órgãos reguladores, estabelece que usar ou implementar, qualquer forma de conexão de compra ou venda de qualquer ativo registrado na Bolsa, adotar dispositivo manipulativo ou enganoso, ou mesmo qualquer artifício em contravenção com as regras e regulações da Securities and Exchange Commission (SEC) prescritas como necessárias ou apropriadas pelo interesse público ou para proteção pelos investidores. [11]

O estatuto da SEC 10b-5 ainda destaca o emprego de dispositivos fraudulentos ou manipulativos, como segue:

É considerado ilegal qualquer pessoa, diretamente ou indiretamente, munido de qualquer subterfúgio que interfira no fluxo de envio de ofertas.

Similar instrução atualizada com a negociação e efetivamente reguladora para o mercado também foi incluído na lei Dodd-Frank (Dodd-Frank Wall Street Reform and Consumer Protection), lei está implementada no ano de 2010 como resposta à crise financeira de 2008. Nomeada desta maneira, devido aos dois principais representantes U.S. Senator Christopher J. Dodd and U.S. Representante Barney Frank. [12], [14], [15]

Deve ser considerado ilegal qualquer pessoa, direta ou indiretamente, no uso ou emprego, ou mesmo tentativa de usar ou empregar, em qualquer conexão de negociação de um instrumento de bolsa, qualquer dispositivo de manipulação ou enganoso ou mesmo artifício, em contravenção de qualquer regra ou regulação de acordo com CFTC (Commodity Futures Trading Commission) promulgada no mínimo 1 ano antes de 21 de julho de 2010. [13], [15]

Em adição estas proibições relacionadas a manipulação, a lei Dodd-Frank adicionou uma orientação específica ao uso de Spoofing e outras práticas utilizadas no mercado de futuros e commodities.

Deve ser considerado ilegal qualquer pessoa que empregue qualquer negociação, prática ou conduta contra as regras:

- (A) Violar as ofertas de compra e venda
- (B) Demonstrar intencional ou imprudente descuido para a execução de negócios durante o período de negociação,
- (C) É caracterizado como, ou comumente conhecido pelos operadores como “Spoofing” (enviar ordens

de compra e vendas com a intenção de cancelar a oferta antes da execução).

A comissão também prove em seu estatuto, quatro possíveis exemplos de situações passíveis na ocorrência de Spoofing, sendo:

- i. Submeter ou cancelar ofertas para sobrecarregar o sistema de cotação
- ii. Submeter ou cancelar ofertas com a intenção de atrasar a execução de outros trades
- iii. Submeter ou cancelar múltiplas ofertas para criar a aparência de uma profundidade de book que não é real.
- iv. Submeter ou cancelar ofertas com a intenção e criar movimentos artificiais de preço de aumento ou redução. [13]

Há alguns trabalhos adotando técnicas diversas para identificar padrões de manipulação no mercado de ações, Zhai no trabalho Computational Intelligent hybrid Model, apresenta um modelo híbrido utilizando SVM (Support Vector Machine) e Hidden Markov Model (HMM) para identificar disruptive trading activities[1], Palshikar e Bahulkar adotaram um modelo baseado lógica temporal fuzzy [22], Sun e seu time propõem uma técnica de multi-slice network analysis[17] e Teema Leangarun, Poja Tangamchit e Suttipong Thajchayapong apresentaram um trabalho de detecção de manipulação de stock price, obtendo bons resultados [26] com a implementação de machine learning e redes neurais.

Outro aspecto importante é sobre a aplicabilidade de trabalhos realizados com base nos dados obtidos em uma bolsa, em outra bolsa do mesmo país ou internacional, devido a especificidade dos dados disponibilizados pelas instituições. Sun e seu time investigaram a identificação da manipulação no mercado de ações da Bolsa da China[23],[24] e Öğüt e seu time fizeram um trabalho de manipulação de mercado no mercado emergente da Istanbul Stock Exchange [25], e os resultados obtidos com as bases delimitadas disponibilizadas pelas instituições, porém estes trabalhos não foram aplicados em outras bolsas para avaliar sua aplicabilidade de forma mais abrangente.

3. ANÁLISE

Os dados separados por instrumento por dia foram compilados respeitando a ordem cronológica de data e tempo, foram selecionados dois instrumentos, um de cada segmento de negociação, DOLZ17, Contrato Futuro de Câmbio de Reais por Dólar Comercial, no mercado de BM&F e PETR4, ação PN da Empresa Petróleo Brasileiro SA para o segmento de Bovespa. Ambos possuem grande volume de negociação e volatilidade em seus respectivos segmentos, de forma a avaliar se a metodologia da análise pode ser compartilhada e principalmente aplicada para ambos os segmentos de negociação, independente da natureza do ativo.

Em uma primeira análise, é possível identificar comportamentos bem diferentes, quanto ao volume de mensagem e a distribuição das mesmas durante o período da amostra do estudo. Esta diferença no comportamento da negociação entre os instrumentos é importante para verificar e atestar a eficiência do método em fluxos de ofertas e características de negociação diferentes.

Em números absolutos, a negociação de PETR4 possui um melhor ratio, ou seja, os investidores precisam de um número menor de ofertas para concretizar um negócio, enquanto os investidores de Dólar, enviam mais ordens para realizar um negócio. O ratio denomina a quantidade de ofertas por negócio, como segue abaixo:

$$\text{ratio} = \left(\sum_{k=0}^m ofc + \sum_{k=0}^n ofv \right) / \sum_{k=0}^p neg$$

Onde:

ofc = somatória da quantidade de ofertas de compra

ofv = somatória da quantidade de ofertas de venda

neg = somatória da quantidade de negócios

A fig. 6 apresenta os dados de negociação dos instrumentos, o ratio maior da Petrobras, demonstra que a negociação desta ação é mais efetiva, ou seja, são enviadas menos ordens para a realização de um negócio.

	PETR4	DOLZ17
Ofertas	4.930.852	7.843.478
Negócios	716.692	674.852
Ratio	0,145	0,086

Fig. 6. Comparação dos dados de negociação dos instrumentos de Petrobras e Futuro de dólar.

Observando os aspectos da negociação de cada instrumento, é identificado a ocorrência de eventos anômalos a negociação, decorrentes do preço inserido com o valor errado, o preço muito acima do praticado no mercado no momento, o maximum price para a ação de Petrobras, é observado o envio sistêmico, ou seja, mesmo preço e horário diariamente, de uma ordem com o preço muito alterado, fora do preço de mercado é inserida pelo mesmo participante. .

O levantamento e estruturação das informações de negociação permitiu observar o padrão de negociação de instrumentos de natureza distintos em segmentos distintos e anormalidades inerentes ao fluxo de cada instrumento em seu mercado.

Está primeira análise superficial denota que a estrutura das informações obtidas a partir dos dados de negociação permite fazer um análise de comportamento dos instrumentos negociados.

Acompanhando o fluxo de mensagens de ambos instrumentos, observamos a diferença da composição dos

books de negociação dos instrumentos (PETR4 e DOLZ17) e a diferença de spreads bem marcante em cada instrumento.

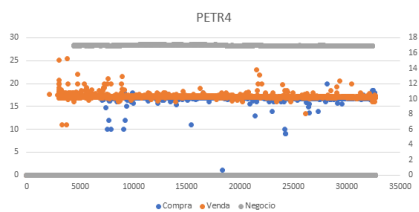


Fig. 8. Apresentação do estudo dos books de compra e venda, em comparação aos negócios realizados de PETR4.

O book de DOLZ17 apresenta uma maior linearidade entre as ordens de compra e venda, e também um menor spread.

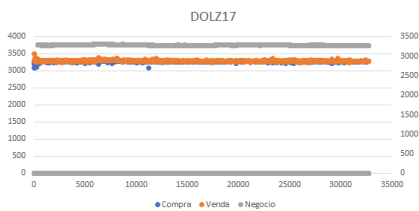


Fig. 9. Apresentação do estudo dos books de compra e venda, em comparação aos negócios realizados de DOLZ17.

O processo de análise, será feito utilizando algoritmos de Machine Learning sem supervisão para identificar os padrões e os relacionamentos entre as informações, e estruturar os dados para que um sistema supervisionado possa rapidamente identificar os padrões de ordem, fig. 10.

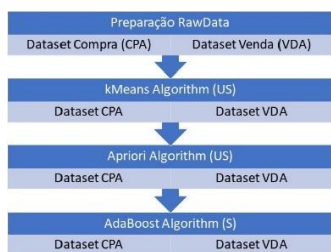


Fig. 10. Apresentação do fluxo de informação entre os algoritmos utilizados na análise.

Fazendo a classificação das ordens de compra, utilizando o dataset de ordens de compra e com a utilização do método de classificação não supervisionado kMeans, observa-se, os grupos de classes de ordem.

O estudo para identificação de ofertas que se enquadrassem no conceito de Spoofing [13], foi feito a partir da análise dos dados dos books de compra e do book de venda de forma de forma separada, pois a interferência no mercado pode ocorrer tanto no lado comprador como lado vendedor.

Com a implementação do método de aprendizagem não supervisionado kMeans, para análise dos dados, foi possível agrupar os dados em clusters. Em uma pré-análise foram identificadas 24 ordens de compra enviadas por uma instituição do mercado financeiro no book de negociação,

que se enquadram na definição do órgão americano [13], pois foram submetidas, possuindo um alto volume financeiro e foram canceladas na sequência, sem a execução de nenhuma ordem.

O estudo listou estas ordens omitindo informações que pudessem facilitar ou mesmo discriminar a identificação da instituição financeira que enviou as ordens, pois são dados reais de negociação e com base no trabalho [1], é necessária uma investigação jurídica para comprovar o real interesse psicológico de manipular a negociação. O trabalho identifica e aponta ordens que devem ser investigadas de maneira mais apurada e possivelmente com dados de fora do sinal de Market data, para comprovar a manipulação dos dados por um processo formal de investigação.

Importante comparar os resultados obtidos com instrumentos líquidos, também com resultados da avaliação de instrumentos de menor liquidez em ambos os mercados. A mesma metodologia foi aplicada a instrumentos menos líquidos do segmento de Bovespa para avaliar a performance do método, os instrumentos AMAR3, BEEF3 e POMO4 foram avaliados. Em paralelo, os instrumentos BGIZ17, BGIX17 e BGIF18, foram avaliados pelo segmento de BMF.

No processo de análise dos grupos de ordens clusterizados, é necessário identificar os clusters de ordens que possuem os menores tempos de permanência (deltatime) das ordens nos respectivos books de compra e venda. Importante a nota que a identificação pelos grupos com menor tempo de permanência no book, ou seja, menor deltatime é decorrente da escolha da manipulação tema deste trabalho, Spoofing e layering, que se utilizam do envio e cancelamento de ordens com o pequeno intervalo de tempo [13], para gerarem liquidez que não existe.

Para atacar este problema foi implementado o algoritmo chamado Apriori Algorithm [42][43], algoritmo de Machine Learning não supervisionado, utilizado em Association Analysis, ou seja, para identificar relacionamentos entre as características de um dataset de maneira autônoma. Um caso clássico de Association Analysis, é o famoso caso de “Beer and Diapers” [45], com uma descrição muito superficial deste famoso caso, foi identificado no estudo que clientes que compravam fraldas também compravam cervejas.

O Apriori Algorithm, foi utilizado nesta etapa da análise para avaliar associação entre as características das ordens. Porém ao invés de identificar as ordens com as características mais representativas, ele foi implementado para identificar características entre as ordens, que menos se repetem, ocorrências isoladas no envio das ordens. Minimizando os valores dos “support” e “confidence”, para que o algoritmo pudesse selecionar as ordens com a maior atipicidade da amostra, e desta forma marcar as ordens como ordens candidatas a avaliação por manipulação do mercado financeiro.

- Support, é definido como a porcentagem do dataset que contém que contém o item analisado [42].
- Confidence, é definido pela regra de associação entre os itens do dataset [42].

Na tabela da fig. 13, são apresentados exemplos de valores de "Support" calculados pelo método Apriori.

Apriori		
BEEF3 CPA - Qty		
Item	Quant.	Support
100	578	0,446332
200	168	0,12973
300	159	0,12278
400	21	0,016216
500	9	0,00695
600	15	0,011583
700	10	0,007722
800	5	0,003861
900	1	0,000772
1000	10	0,007722
1100	2	0,001544
1200	1	0,000772
1300	2	0,001544
1500	1	0,000772
1600	1	0,000772
1700	1	0,000772
1800	1	0,000772
1900	2	0,001544
2000	1	0,000772
2400	2	0,001544
4200	1	0,000772
4300	27	0,020849
4400	201	0,155212
4500	71	0,054826
5000	4	0,003089
9000	1	0,000772

Fig. 13. Exemplo dos valores de support calculado pelo Apriori. * a tabela não possui todos os valores, em virtude do tamanho da tabela, os primeiros valores foram selecionados para exemplificação

Na Fig 14. a representação gráfica dos "Support" de três itens calculados para o book de compra do instrumento BEEF3, um instrumento com liquidez menor, em comparação a liquidez dos instrumentos de PETR4 (Bovespa) e DOLZ17 (BMF), segundo o volume de mensagens.



Fig. 14. Exemplo dos valores de support calculados para o book de compra do instrumento BEEF3.

Com o resultado dos "Supports", calculados com os itens que possuíam um valor igual ou menor a 0,01, contemplaram o cálculo do grau de confiança "confidence", as ordens foram selecionadas como candidatas recebendo uma identificação positiva no dataset (1) na avaliação de manipulação do mercado no dataset.

Ao término desta etapa é possível listar as ordens do cluster que possuem características suspeitas e que demandem uma investigação mais detalhada com uma interação humana sobre as motivações do envio da ordem pela instituição no mercado [1].

Como última etapa da avaliação das ordens, foi implementado o algoritmo de Adaboost, tendo como input de treinamento e teste, para identificação em tempo real o resultado da análise de associação feita via Apriori Algorithm. Desta forma os padrões identificados nas ordens podem rapidamente ser identificados e reportados como pontos de análise para avaliação detalhada por humanos.

O adaboost é um algoritmo de machine learning supervisionado [46], e recebe como fonte para análise de padrão o resultado obtido da análise se associação feita pelo Apriori. Desta forma é possível buscar a identificação dos padrões aprendidos durante análise nas ordens em tempo real.

Um problema importante que se materializa neste tipo de análise de manipulação é o desbalanceamento das classes[42][46], pois a grande maioria dos casos de ordens no estão no mercado, foram enviadas corretamente com a intenção de negociar, e apenas pouquíssimas ordens são sinalizadas com características de manipulação ou mesmo falhas operacionais, sendo identificadas como ocorrências[16]. Na figura 15, exemplos de alguns instrumentos de Bovespa e BMF, com a quantidade de ocorrências. É possível notar o desbalanceamento de ocorrências.

AMAR3					
CPA			VDA		
Classific.	Total	%	Classific.	Total	%
1	28	0,0502693	1	24	0,027874564
-1	529	0,9497307	-1	837	0,972125436

POMO4					
CPA			VDA		
Classific.	Total	%	Classific.	Total	%
1	50	0,045620438	1	16	0,013816926
-1	1046	0,954379562	-1	1142	0,986183074

Fig. 15. Exemplo dos valores de support calculados para o book de compra do instrumento BEEF3.

Para atacar este problema foi feito um dataset a partir dos dados avaliados, utilizando resampling, ou seja, preparando um dataset para a avaliação pelo Adaboost, com a reinserção de ordens selecionadas com características de manipulação, junto com as ordens avaliadas como normais, sendo escolhidas de forma aleatória no conjunto de dados. Esta técnica foi adotada para minimizar a possibilidade de um aprendizado com viés errôneo pelo algoritmo, ou mesmo um overfit da análise de dados, no quadro abaixo, a composição dos dataset já balanceados.

AMAR3					
CPA			VDA		
Classific.	Total	%	Classific.	Total	%
1	600	0,6	1	600	0,6
-1	400	0,4	-1	400	0,4

POMO4					
CPA			VDA		
Classific.	Total	%	Classific.	Total	%
1	656	0,599634369	1	694	0,599827139
-1	438	0,400365631	-1	463	0,400172861

Fig. 16. Composição do dataset de treinamento do Adaboost.

Instrumentos pouco líquido que apresentaram menos de 1 mil ordens dia, tiveram um dataset limitado 1 mil inserções no dataset, os instrumentos negociados com o volume até 10000 ordens dias, mantiveram a quantidade de ordens no dataset. O instrumento POMO4, na figura 16, apresentou 1094 ordens de compra e 1157 ordens de venda, desta forma o dataset compreendeu a mesma quantidade de inserções em seu respectivo dataset. Os instrumentos mais líquidos com mais de 10000 ordens dias, mantiveram o dataset limitado a 10000 inserções do dataset.

Esta limitação foi implementada pois apesar da grande quantidade de ordens enviadas no dia o número de ordens selecionadas como suspeitas não aumenta proporcionalmente, e a reinserção das ordens pela técnica resampling, força uma situação de overfitting.

A utilização do Adaboost, permite avaliar o comportamento das ordens com base nas características enviadas no book de negociação. As ordens com perfil mais longo como Good till Cancel (GTC), Good till Date(GTD) e Day de permanecerem no book[47], terão um delay no início da avaliação, ao passo que ordens que forem inseridas e canceladas na sequencia serão analisadas na sequencia dos fatos, dando a monitoração a agilidade da comunicação junto aos executores das ordens.

Por fim, como resultado do trabalho um quadro resumo fig. 17, com algumas ordens de instrumentos de ambos os segmentos BEEF3 (Segmento Bovespa, menor liquidez) e DOLZ17 (Segmento BM&F, muito líquido), ilustrando a capacidade do método de identificar padrões anômalos para instrumentos distintos, independente do comportamento de negociação dos mesmos.

A ordem selecionada em verde, foi considerada pelo método, como uma ordem ok, foi adicionada ao quadro, como comparação em relação as demais ordens selecionadas com movimentação anômala, “suspeita” e demandam uma análise humana [1]. Para as ordens de DOLZ17, foram adicionadas as ordens identificadas como “suspeitas” apenas, considerando características diversas das ordens. Informações que pudessem identificar o código da corretora ou mesmo o código da ordem (secondary order id) [47], foram omitidos para manter a confidencialidade dos responsáveis pela ordem.

Data	Instrumento	B/S	Cód Ordem	Atualizações	Preço Fim	Quantidade	Código Corretora	Tempo da Ordem	Preço Negócio
2017-11-03	BEEF3	CPA	Ordem A	47	11.14	5000	Corretora A	0:09:39.718000	11.26
2017-11-03	BEEF3	CPA	Ordem B	15	11.26	5000	Corretora A	0:01:12.863000	11.37
2017-11-03	BEEF3	CPA	Ordem C	22	11.09	5000	Corretora A	0:00:59.387000	11.21
2017-11-03	BEEF3	CPA	Ordem D	2	11.19	2400	Corretora B	0:00:00	11.2
2017-11-03	BEEF3	CPA	Ordem E	2	11.21	2400	Corretora B	0:00:00	11.21
2017-11-03	DOLZ17	VDA	Ordem F	3480	3321.0	50	Corretora C	0:00:13.048000	3316.5
2017-11-03	DOLZ17	VDA	Ordem G	5	3315.5	770	Corretora D	0:00:02.289000	3314.0

Fig. 17. Seleção de ordens classificadas com movimentação suspeita.

4. CONCLUSÃO

A metodologia proposta neste trabalho não é absoluta em identificar todo e qualquer processo de manipulação no mercado financeiro de ações. Neste trabalho foram avaliados apenas o Spoofing e Layering por ativo.

Foram utilizadas técnicas de algoritmos sem supervisão como kMeans e Apriori, e algoritmo supervisionado Adaboost, trabalhando de maneiras complementares na identificação de padrões, mapeamento de relevância dos mesmos e a rápida identificação dos mesmos através de algoritmos supervisionados.

Os dados de Market data são muito ricos em detalhe de informação, a partir destes é possível reproduzir todas as movimentações do book de ações de todos os ativos negociados, permitindo trabalhar com agilidade nas movimentações que apresentem distorções em relação as demais.

Importante destacar, no entanto, apesar dos detalhes de negociação o Market data não discrimina o autor da ordem de maneira publica, apesar de possuir identificadores em sua mensageria que permitam aos responsáveis por suas ordens identifica-las. Desta forma a metodologia não identifica os autores do envio das ordens, mas sim seleciona as ordens, instituições e principalmente a situação pela qual a ordem foi escolhida.

A metodologia adotada, apresentou uma abordagem técnica somando algoritmos sem supervisão com algoritmo supervisionado, de forma que algoritmos sem supervisão, identificando padrões atípicos a negociação, montassem uma base de aprendizado para que um algoritmo supervisionado identificasse padrões com agilidade de forma satisfatória, porém suscetível a liquidez, ou melhor discriminando, volume de atualização das ordens de cada instrumento. A aplicação da Lei Dodd-Frank [13], escolhida como norte regulatório deste estudo técnico no Market data brasileiro, não há torna uma referência para todo e qualquer mercado, podendo esta metodologia se adaptar a outras regulações legais relacionadas ao mercado acionário.

Referências

- [1] Decision Support Systems; Computational intelligent hybrid model for detecting disruptive trading activity, Zhai, Jiz ; Cao, Yi ; Ding, Xuemei, Ding ; Yuhua, Li ; ELSIVIER ; Journal
- [2] BLOOMBERG, Bloomberg do Brasil, 2017 em : [https://www.bloomberg.com/news/articles/2017-05-03/exchanges-still-dreaming-of-deals-after-lse-deutsche-boerse-dies - Two Market Giants Tried to Create a \\$75 Billion Mega-Exchange, Matthew Leising, 3 de maio de 2017 06:00 BRT Updated on 3 de maio de 2017 11:29 BRT. Acesso em: 13/05/2018](https://www.bloomberg.com/news/articles/2017-05-03/exchanges-still-dreaming-of-deals-after-lse-deutsche-boerse-dies - Two Market Giants Tried to Create a $75 Billion Mega-Exchange, Matthew Leising, 3 de maio de 2017 06:00 BRT Updated on 3 de maio de 2017 11:29 BRT. Acesso em: 13/05/2018)
- [3] B3, Market Data: Histórico Pessoas Físicas. Financeiro, São Paulo, 01/09/2017. Disponível em: http://www.bmfbovespa.com.br/pt_br/servicos/market-data/consultas/historico-pessoas-fisicas . Acesso em: 05/08/2018.

- [4] VALOR, Valor Econômico, disponível em: <http://www.valor.com.br/valor-investe/casa-das-caldeiras/5058252/os-donos-da-bolsa>. Angela Bittencourt, 28/07/2017 às 11h19. Acesso em: 13/05/2018
- [5] VALOR, Valor Econômico, disponível em: <http://www.valor.com.br/financas/3025114/bmfbovespa-inicia-migracao-de-negocios-com-aco-es-para-plataforma-puma>, Vinicius Pinheiro 27/02/2013 às 14h53. Acesso em: 13/05/2018.
- [6] B3, PUMA Trading System BM&FBOVESPA, disponível em: http://www.bmfbovespa.com.br/pt_br/servicos/negociacao/puma-trading-system-bm-fbovespa/. Acesso em: 15/05/2018
- [7] EXAME, disponível em: <https://exame.abril.com.br/mercados/bovespa-estrea-sistema-mais-rapido-que-1-milésimo-de-segundo/> - Bovespa estreia sistema mais rápido que 1 milésimo de segundo, Marcel Salim, 24/08/2011 às 22h56. Acesso em 13/05/2018.
- [8] ESMA, European Securities and Markets Authority, <https://www.esma.europa.eu/about-esma/who-we-are>, ESMA. Acesso em 13/05/2018
- [9] ESMA, MIFID II disponível em: <https://www.esma.europa.eu/policy-rules/mifid-ii-and-mifir>, ESMA. Acesso 13/05/2018
- [10] BSM, Instrução CVM 8, <http://www.cvm.gov.br/legislacao/instrucoes/inst008.html>, CVM. Acesso 13/05/2018.
- [11] SEC, US Securities And Exchange Commission, <https://www.sec.gov/answers/about-lawsshtml.html#secexact1934>, The Laws That Govern the Securities Industry Securities Exchange Act of 1934
- [12] CFTC, Antidruptive Practices Authority No. 3038-AD96, Commodity Exchange Act Release, May 20, 2013.
- [13] Dodd-Frank Wall Street Reform and Consumer Protection Act § 753 (Anti-Manipulation Authority) 7 U.S.C. § 9.1 (2016).
- [14] INVESTOPEDIA, <https://www.investopedia.com/terms/d/dodd-frank-financial-regulatory-reform-bill.asp>, Dodd-Frank Wall Street Reform and Consumer Protection Act, Acesso em 13/05/2018.0
- [15] Montgomery, John, Spoofing, Market Manipulation, and the Limit-Order Book, May 3, 2016, Navigant
- [16] Hafiz, Kosemani Temitayo, Aghili, Dr. Shaun, Zavarsky, Dr. Pavol; The use of Predictive Analytics Technology to Detect Credit Card Fraud in Canada, Department of Information Systems Security and Assurance management. Concordia University of Edmonton; Edmonton, Canada
- [17] Sun, Xiao-Qian; Shen, Hua-Wei; Cheng, Xue-Qi; Zhang, Yuqing; Detecting anomalous traders using multi-slice network analysis; ELSEVIER; Journal
- [18] F. Allen, D. Gale, Stock-price manipulation, Rev. Financ. Stud. 5 (3) (1992) 503–529.
- [19] R.A. Jarrow, Market manipulation, bubbles, corners, and short squeezes, J. Financ. Quant. Anal. 27 (03) (1992) 311–336.
- [20] X. Sun, X. Cheng, H. Shen, Z. Wang, Statistical properties of trading activity in Chinese stock market, Phys. Procedia 3 (5) (2010) 1699–1706.
- [21] R.A. Jarrow, Market manipulation, bubbles, corners, and short squeezes, J. Financ. Quant. Anal. 27 (03) (1992) 311–336.
- [22] G.K. Palshikar, A. Bahulkar, Fuzzy temporal patterns for analysing stock market databases, in: Proceedings of the International Conference on Advances in Data management, (COMAD), Tata-McGraw Hill, Pune, India, 2000, pp. 135–142.
- [23] X.-Q. Sun, X.-Q. Cheng, H.-W. Shen, Z.-Y. Wang, Distinguishing manipulated stocks via trading network analysis, Physic A 390 (20) (2011) 3427–3434.
- [24] X.-Q. Sun, H.-W. Shen, X.-Q. Cheng, Z.-Y. Wang, Degree-strength correlation reveals anomalous trading behavior, PLoS One 7 (10) (2012) e45598.
- [25] H. Ögüt, M.M. Doganay, R. Aktas, Detecting stock-price manipulation in an emerging market: The case of turkey, Expert Syst. Appl. 36 (9) (2009) 11944–11949.
- [26] Leangarun, Teema; Tangamchit Poj; Thachayapong, Suttipong; Stock Price Manipulation Detection using a computational Neural Network Model, 2016
- [27] Fonooni, Benjamin; Moghadam, S. J. Mousavi; Applying induced Aggregation Operator in Designing Intelligent Monitoring System for Financial Market.
- [28] Ai, Xiao-Wei; Hu, Tianming; Li, Xi; Xiong, Hui; Clustering High-frequency Stock Data for Trading Volatility Analysis; 2010; Ninth International Conference on Machine Learning and Applications.
- [29] Tsantekidis, Avraam; Passalis; Nikolaos; Tefas, Anastasios; Kannianen, Juho; Gabbou, Moncef; Iosifidis, Alexandros; Forecasting Stock prices from limit order book using Convolutional neural networks.
- [30] D.J. Cumming, F. Zhan, M.J. Aitken, High frequency trading and end-of-day price dislocation, October 28, 2013 Retrieved from Social Science Research Network, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2145565.

- [31] F. Allen, D. Gale, Stock price manipulation, *The Review of Financial Studies* 5 (3) (1992) 503–529.
- [32] Y. Cao, Li, Coleman, Belatreche, McGinnity, Detecting wash trade in financial market using digraphs and dynamic programming, *IEEE Transactions on Neural Networks and Learning Systems* (2015), <http://dx.doi.org/10.1109/TNNLS.2015.2480959>.
- [33] Y. Cao, Y. Li, S. Coleman, A. Belatreche, M. McGinnity, Adaptive hidden Markov model with anomaly states for price manipulation detection
- [34] R. Ghazali, A.J. Hussain, N.M. Nawawi, B. Mohamad, Non-stationary and stationary prediction of financial time series using dynamic ridge polynomial neural network, *Neurocomputing* 72 (10–12) (2009) 2359–2367.
- [35] KELLER, John D. ; NAMEE, Brian Mac ; ARCY, Aoife D' ; *Fundamentals of Machine Learning for Predictive Data Analytics Algorithms, Worked Examples and case studies*. MIT Press, Cambridge Massachusetts
- [36] Equity Market Structure Literature Review, Part II High Frequency Trading by Staff of the Division of Trading and Markets, 03/18/2014, Staff of U.S. Securities and Exchange Commission.
- [37] CME, Rule 575, August 2014 Retrieved from U.S. Commodity Futures Trading Commission, <http://www.cftc.gov/filings/orgrules/rule082814cmcdcm001.pdf>.
- [38] Wang, X ; Wellman P. ; Michael, Spoofing the Limit Order Book : An Agent-Based Model ; *Computer Science & Engineering*, University of Michigan ; Michigan
- [39] E. Martinez-Miranda, P. McBurney, and M. Howard. ; Learning unfair trading: A market manipulation analysis from the reinforcement learning perspective. ; In *IEEE International Conference on Evolving and Adaptive Intelligent Systems*, pages 103–109, 2016.
- [40] Lee, J. Eun ; Eom, S. Kyong ; Park, K. Suh ; Microstructure-based manipulation : Strategic behavior and performance of spoofing traders ; *ELSEVIER ; Journal of Financial Markets* 16 (2013), pages 227-252 ; 2012
- [41] Market Data BMFBOVESPA: <ftp://ftp.bmf.com.br/MarketData/>, Origem dos Dados.
- [42] Harrington, Peter; *Machine Learning in Action*; Manning; Manning Publications Co.; United States, 2012
- [43] YANG Q. FU, C. W. Q.; YANG, J. A matrix-based apriori algorithm improvement. 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), n. 1, p. 824–828, 2018.
- [45] Cohen, N., Data Mining: Naggging that it really adds up, 2000, URL http://www.open-mag.com/features/Vol_16/datamining/datamining.htm
- [46] Schapire, Robert E.; Freund, Yoav; *Boosting Foundations and Algorithms*; MIT Press; United States; 2014
- [47] B3, http://www.b3.com.br/pt_br/solucoes/plataformas/puma-trading-system/para-desenvolvedores-e-vendors/entrypoint-entrada-de-ofertas/; EntryPoint; EntryPoint: Interface de Envio de Ordens - Regras - versão 2.9.5 (em Inglês); 28/12/2018