



PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

A importância da inteligência artificial para a liquidez do Mercado de Capitais Brasileiro.

Lucas Dessanti Rebehy

TRABALHO DE CONCLUSÃO DE CURSO

CENTRO DE CIÊNCIAS SOCIAIS - CCS

DEPARTAMENTO DE ADMINISTRAÇÃO

Graduação em Administração de Empresas

Rio de Janeiro, Novembro de 2018.



Lucas Dessanti Rebehly

**A importância da inteligência artificial para a liquidez do
Mercado de Capitais Brasileiro.**

Trabalho de Conclusão de Curso

Trabalho de Conclusão de Curso, apresentado ao programa de graduação em Administração da PUC-Rio como requisito parcial para a obtenção do título de graduação em Administração.

Orientador: Antonio Carlos Figueiredo

Rio de Janeiro, Novembro de 2018.

Agradecimentos

Ao professor António Carlos Figueiredo, por me orientar sobre quais eram os melhores caminhos a seguir, tirando dúvidas e acreditando na proposta que pretendo entregar neste trabalho de conclusão de curso.

Aos meus avós, Lely Said Chafic Rebehy e Maria Eliana Alves Rebehy, por acreditarem em minha capacidade e sempre investirem em minha formação acadêmica.

Aos meus pais, Lely Rebehy Júnior e Giuliana Rita Desanti, e ao meu irmão, Henrique, por contribuírem com minha educação.

Finalmente, gostaria de agradecer a todos os meus amigos que permaneceram comigo durante todo o curso e aos demais excelentes professores que compõem o quadro da Puc - RJ.

Resumo

Rebehy, Lucas Dessanti. A importância da inteligência artificial para a liquidez do Mercado de Capitais Brasileiro. Rio de Janeiro, 2018. Número de páginas 32. Trabalho de Conclusão de Curso – Departamento de Administração. Pontifícia Universidade Católica do Rio de Janeiro.

O mercado financeiro e de capitais vem rompendo paradigmas desde 2008 após a ruptura do pregão viva-voz para o eletrônico. O objetivo deste artigo é discutir a importância da inteligência artificial para a liquidez do Mercado de Capitais Brasileiro. Os argumentos usados nos mostra que a existência de alternativas tecnológicas voltadas para o HFT (high frequency trading), sigla em inglês pela qual o sistema é conhecido e AlgoTradings são essenciais para o desenvolvimento dessa indústria. Ao final realizamos um teste que confirma a hipótese de que o volume operado vem aumentando X Bid-Ask spread vem diminuindo, indicando uma evolução na liquidez.

Palavras-chave

Inteligencia Artificial, Machine Learning, Big Data, Algoritmos, HFT, DMA, Bid-Ask spread, Liquidez

Abstract

Rebehy, Lucas Dessanti. The importance of artificial intelligence for the liquidity of the Brazilian Capital Market. Number of pages 32. Course Completion Work - Administration Department. Pontifical Catholic University of Rio de Janeiro.

The financial and capital markets have been breaking paradigms since 2008 after the breakup of the hands-off to electronic trading. The purpose of this article is to discuss the importance of artificial intelligence for the liquidity of the Brazilian Capital Market. The arguments used show us that the existence of technological alternatives focused on the HFT (high frequency trading), known in English and AlgoTradings are essential for the development of this industry. At the end we perform a test that confirms the hypothesis that the volume operated has been increasing X the Bid-Ask spread has been decreasing, indicating an evolution in liquidity.

Key-words

Artificial Intelligence, Machine Learning, Algorithms, HFT, DMA, Bid-Ask spread, Liquidity

Sumário

| | | |
|------|---|----|
| 1. | O tema e o problema de estudo | 8 |
| 1.1. | Introdução ao tema e ao problema do estudo..... | 8 |
| 1.2. | Objetivo do estudo | 9 |
| 1.3. | Objetivos intermediários do estudo | 9 |
| 1.4. | Delimitação e foco do estudo | 9 |
| 1.5. | Relevância do Estudo | 9 |
| 2. | Contexto e realidade investigada | 10 |
| 2.1. | Inteligência Artificial | 10 |
| 2.2. | Aprendizado de máquina (machine learning)..... | 11 |
| 2.3. | Big Data..... | 16 |
| 2.4. | Algoritmos de Trading (AlgoTrading) | 17 |
| 2.5. | Direct Market Access (DMA)..... | 19 |
| 2.6. | High Frequency Trading (HFT) | 24 |
| 3. | Métodos e procedimentos de coleta e de análise de dados do estudo | 27 |
| 3.1. | Estratégia de investigação OU Métodos de pesquisa utilizados | 27 |
| 3.2. | Procedimentos e instrumentos de coleta de dados utilizados no estudo..... | 27 |
| 3.3. | Formas de tratamento e análise dos dados coletados para o estudo..... | 28 |
| 3.4. | Limitações do método..... | 29 |
| 4. | Apresentação e análise dos resultados | 30 |
| 4.1 | Análise do Average Bid-Ask spread..... | 30 |
| 4.2 | Análise da evolução do Volume..... | 31 |
| 5. | Conclusões e recomendações para novos estudos..... | 32 |
| 6. | Referências Bibliográficas | 33 |

Lista de figuras

| | |
|---|----|
| Figura 1-1: Modelo 1 – DMA Tradicional | 20 |
| Figura 1-2: Modelo 2 – DMA via Provedor | 21 |
| Figura 1-3: Modelo 3 – DMA via Conexão Direta | 21 |
| Figura 1-4: Modelo 4 –DMA via Conexão Direta Co-location | 23 |
| Figura 2: Função de densidade de probabilidade para a distribuição normal.. | 29 |

Lista de Tabela

| | |
|--|----|
| Tabela 1.1: Teste Z para o Average Bid-Ask spread..... | 30 |
| Tabela 1.2 – Teste Z para a evolução do Volume | 31 |

1. O tema e o problema de estudo

1.1. Introdução ao tema e ao problema do estudo

Que papel a inteligência artificial assume na liquidez do Mercado de Capitais Brasileiro?

O mercado financeiro e de capitais vem rompendo paradigmas desde 2008 após a ruptura do pregão viva-voz para o eletrônico. Tal ruptura só se pode ser sentida no Brasil no final de 2011, quando a modalidade DMA 3 e 4 foram oferecidas pela antiga BM&FBovespa. O objetivo deste artigo é discutir a importância da inteligência artificial para a liquidez do Mercado de Capitais Brasileiro. Os argumentos usados nos mostra que a existência de alternativas tecnológicas voltadas para o HFT (high frequency trading), sigla em inglês pela qual o sistema é conhecido e AlgoTradings (Algoritmos que executam ordens de compra e venda automaticamente) são essenciais para o desenvolvimento dessa indústria.

Além desta breve introdução, o artigo conta com outras quatro seções. A segunda explica o contexto da inteligência artificial, bem como a sua relação com Big data e Algoritmos, muito usados não só para processos internos como para execuções de ordens em alta frequência (HFT).

A terceira seção levantará os Métodos e procedimentos de coleta e de análise de dados do estudo que tem como base dois principais indicadores, volume e Bid-Ask spread.

A quarta seção apresenta o resultado com base em teste estatístico da evolução no contexto brasileiro. E identifica como a inteligência artificial influencia a liquidez do Mercado de Capitais Brasileiro. E, finalmente, a quinta apresenta algumas conclusões sobre o estudo.

1.2. Objetivo do estudo

O objetivo deste artigo é discutir e identificar por meio de teste estatístico a importância que a inteligência artificial assume para a liquidez do Mercado de Capitais Brasileiro.

1.3. Objetivos intermediários do estudo

Conceituar a Inteligência artificial, o aprendizado de máquina (em inglês, machine learning) e como eles se relacionam com o Big Data e os Algoritmos. levantando como é fundamental a inteligência artificial para a liquidez do Mercado de Capitais Brasileiro.

1.4. Delimitação e foco do estudo

Contextualizar Inteligência artificial, aprendizado de máquina (machine learning), Big Data, Algoritmos de Trading (AlgoTrading) e os 4 módulos de DMA (Direct Market Access) e High Frequency Trading (HFT). Além disso será feito um teste estatístico com base no Volume de ações que compõem o índice IBOVESPA e a média diária do Bid-Ask spread das respectivas ações.

1.5. Relevância do Estudo

Por ser um conceito muito novo e ainda com dificuldades de ser sintetizado, o objeto dessa pesquisa veio para ficar e rompeu barreiras jamais imaginadas no mercado financeiro e de capitais. Empresas e gestores poderão se beneficiar e agregar valor a si mesmos ao ler esse projeto final de conclusão de curso. Amantes do assunto, porém ainda leigos, terão o conforto de ler textos não obscuros, de fácil compreensão.

2. Contexto e realidade investigada

Neste capítulo são apresentados e discutidos aspectos conceituais e estudos relacionados ao tema e estudo em investigação e que servirão de base para a análise realizada.

2.1. Inteligência Artificial

Os entendidos em ficção científica acreditam que Inteligência Artificial é sobre encontrar padrões, aprendendo com os erros, dividindo problemas, e adaptando-se para conseguir objetivos específicos. Mas na verdade IA é uma série de ferramentas de lógica incríveis que permitem progressão metodológica no processamento de dados. Parece complexo, mas quando olhamos para seus componentes-base torna-se fácil de entender.

A idéia de computadores superando e substituindo seres humanos já existe em filmes e livros há décadas. Felizmente, isso ainda não aconteceu em grande escala. Mas o que aconteceu é o surgimento recente de conceitos de inteligência artificial - especificamente a computação cognitiva.

Esses conceitos envolvem plataformas de tecnologia avançada que podem abordar situações complexas caracterizadas por ambiguidades e incertezas. A computação cognitiva começou a aumentar o poder das decisões empresariais e o desempenho de computadores ao lado do processo de pensamento humano e da análise tradicional. De fato, o domínio do gerenciamento de riscos se presta particularmente bem às capacidades de computação cognitiva, pois as questões típicas de risco geralmente incluem eventos improváveis e / ou ambíguos. As empresas e as organizações do setor financeiro progrediram em termos de uso de quantidades maciças de dados internos e externos para assumirem uma posição de risco mais preventiva, diz Samir Hans, diretor-presidente de Deloitte na prática de

Forense e Investigações da Deloitte Transactions e Business Analytics LLP. No entanto, os métodos tradicionais de análise tornaram-se cada vez mais incapazes de lidar com esse volume de dados. Em vez disso, as capacidades cognitivas - incluindo a mineração de dados, a aprendizagem de máquinas e o processamento de linguagem natural - superam as análises tradicionais e são aplicadas contra esses conjuntos maciços de dados para ajudar a encontrar indicadores de riscos conhecidos e desconhecidos, podendo assim converter tais indicadores em tomadas de decisões estratégicas.

O termo inteligência artificial foi criado em 1956, mas só se popularizou hoje graças aos crescentes volumes de dados disponíveis, algoritmos avançados, e melhorias no poder e no armazenamento computacionais.

As primeiras pesquisas de IA nos anos 1950 exploraram temas como a resolução de problemas e métodos simbólicos. Na década de 1960, o Departamento de Defesa dos EUA se interessou por este tipo de tecnologia e começou a treinar computadores para imitar o raciocínio humano básico. Por exemplo, a Defense Advanced Research Projects Agency (DARPA) completou um projeto de mapeamento de ruas nos anos 1970. E a DARPA criou assistentes pessoais inteligentes em 2003, muito tempo antes de Siri, Alexa ou Cortana serem nomes comuns do nosso cotidiano.

Esses primeiros trabalhos prepararam o caminho para a automação e o raciocínio formal que vemos nos computadores de hoje, incluindo sistemas de apoio à decisão e sistemas inteligentes de pesquisa que podem ser projetados para complementar e expandir as capacidades humanas.

Enquanto os filmes de Hollywood e os romances de ficção científica retratam a inteligência artificial como robôs humanoides que dominam o mundo, a evolução atual das tecnologias de IA não é tão assustadora – ou tão inteligente assim. Em vez disso, a IA evoluiu para fornecer muitos benefícios específicos para todas as indústrias.

2.2. Aprendizado de máquina (machine learning)

O aprendizado de máquina (em inglês, machine learning) é um método de análise de dados que automatiza a construção de modelos analíticos. É um ramo da inteligência artificial baseado na ideia de que sistemas podem aprender com dados, identificar padrões e tomar decisões com o mínimo de intervenção humana.

Graças às novas tecnologias computacionais, o machine learning de hoje não é como o machine learning do passado. Ele nasceu do reconhecimento de padrões e da teoria de que computadores podem aprender sem serem programados para realizar tarefas específicas; pesquisadores interessados em inteligência artificial queriam saber se as máquinas poderiam aprender com dados. O aspecto iterativo do aprendizado de máquina é importante porque, quando os modelos são expostos a novos dados, eles são capazes de se adaptar independentemente. Eles aprendem com computações anteriores para produzir decisões e resultados confiáveis, passíveis de repetição. Isso não é uma ciência nova – mas uma ciência que está ganhando um novo impulso.

Enquanto a inteligência artificial (IA) pode ser definida, de modo amplo, como a ciência capaz de mimetizar as habilidades humanas, o machine learning é uma vertente específica da IA que treina máquinas para aprender com dados.

O interesse renovado no aprendizado de máquina se deve aos mesmos fatores que tornaram a mineração de dados e a análise Bayesiana mais populares do que nunca: coisas como os crescentes volume e variedade de dados disponíveis, o processamento computacional mais barato e poderoso, o armazenamento de dados acessível etc.

Tudo isso significa que é possível produzir, rápida e automaticamente, modelos capazes de analisar dados maiores e mais complexos, e entregar resultados mais rápidos e precisos – mesmo em grande escala. E ao construir modelos precisos, uma organização tem mais chances de identificar oportunidades lucrativas – ou de evitar riscos desconhecidos.

O que é preciso para criar bons sistemas de machine learning?

- Capacidade de data preparation;
- Algoritmos – básicos e avançados;
- Processos automatizados e iterativos;
- Escalabilidade;
- Modelagem conjunta.

Machine learning não é uma tecnologia específica em si; ele envolve softwares como de mineração de dados e advanced analytics para processar grandes volumes de dados e descobrir insights.

A maioria das indústrias que trabalham com grandes quantidades de dados tem reconhecido o valor da tecnologia de aprendizado de máquina. Ao extrair insights desses dados – frequentemente em tempo real – as organizações são capazes de trabalhar com mais eficiência ou de ganhar uma vantagem competitiva sobre seus concorrentes.

Bancos e outros negócios na indústria financeira usam tecnologias de machine learning para dois propósitos principais: identificar insights importantes nos dados e prevenir fraudes. Esses insights podem identificar oportunidades de investimento, ou ajudar investidores a saber quando fazer o trade. A mineração de dados também pode identificar clientes com alto perfil de risco ou usar cyber-vigilância para encontrar sinais de fraudes.

Dois dos métodos mais adotados de machine learning são o **aprendizado supervisionado** e o **aprendizado não-supervisionado** –

Algoritmos de aprendizado **supervisionado** são treinados por meio de exemplos rotulados, como uma entrada na qual a saída desejada é conhecida. Por exemplo, um equipamento poderia ter pontos de dados rotulados como “F” (falha) ou “E” (executa). O algoritmo de aprendizado recebe um conjunto de entradas junto com as saídas corretas correspondentes, e aprende ao comparar a saída real com as saídas corretas para encontrar erros. Ele, então, modifica o modelo de acordo. Através de métodos como classificação, regressão e gradient boosting, o aprendizado supervisionado utiliza padrões para prever os valores de rótulos em dados não-rotulados adicionais. O aprendizado

supervisionado é comumente empregado em aplicações nas quais dados históricos preveem eventos futuros prováveis. Por exemplo, ele pode antecipar quando transações via cartão de crédito são passíveis de fraude ou qual segurado tende a reivindicar sua apólice.

O aprendizado **não-supervisionado** é utilizado contra dados que não possuem rótulos históricos. A "resposta certa" não é informada ao sistema. O algoritmo deve descobrir o que está sendo mostrado. O objetivo é explorar os dados e encontrar alguma estrutura dentro deles. O aprendizado não-supervisionado funciona bem com dados transacionais. Por exemplo, ele pode identificar segmentos de clientes com atributos similares que podem, então, ser tratados de modo igualmente similar em campanhas de marketing; ou ele pode encontrar os principais atributos que separam segmentos distintos de clientes. Técnicas populares incluem mapas auto-organizáveis, mapeamento por proximidade, agrupamento k-means e decomposição em valores singulares. Esses algoritmos também são utilizados para segmentar tópicos de texto, recomendar itens e identificar pontos discrepantes nos dados.

2.3. Big Data

Big data é um termo que descreve o grande volume de dados — tanto estruturados quanto não-estruturados — que sobrecarrega as empresas diariamente. Mas não é a quantidade de dados disponíveis que importa; é o que as organizações fazem com eles. O big data pode ser analisado para obter insights que levam a decisões melhores e ações estratégicas de negócio.

Embora o termo "big data" seja relativamente novo, o ato de coletar e armazenar grandes quantidades de informações para análises eventuais é muito antigo. O conceito ganhou força no começo dos anos 2000, quando o analista Doug Laney articulou a definição atualmente mainstream de big data em três Vs:

Volume. Organizações coletam dados de fontes variadas, incluindo transações financeiras, mídias sociais e informações de sensores ou dados transmitidos de máquina para máquina. No passado, armazená-los teria sido um problema — mas novas tecnologias (como o Hadoop) aliviaram esse fardo.

Velocidade. Os dados são transmitidos numa velocidade sem precedentes e devem ser tratados em tempo hábil. Etiquetas RFID, sensores e medições inteligentes estão impulsionando a necessidade de lidar com torrentes de dados praticamente em tempo real.

Variedade. Dados são gerados em inúmeros formatos — desde estruturados (numéricos, em databases tradicionais) a não-estruturados (documentos de texto, e-mail, vídeo, áudio, cotações da bolsa e transações financeiras).

A quantidade de dados criados e armazenados globalmente é quase inimaginável, e apenas continua crescendo. Isso significa que há ainda mais potencial para extrair insights importantes dessas informações — embora apenas uma pequena porcentagem dos dados seja realmente analisada. O que isso significa para as empresas? Como elas podem fazer um uso melhor dessas informações brutas que fluem para dentro de suas paredes todos os dias?

A importância do big data não gira em torno da quantidade de dados que você tem, mas do que você faz com eles. Você pode obter dados de várias fontes e analisá-los para encontrar respostas que permitem 1) reduzir custos; 2) economizar tempo; 3) desenvolver novos produtos e otimizar ofertas; 4) tomar decisões mais inteligentes. Quando você combina big data com inteligência analítica de alta performance, você pode realizar tarefas corporativas como:

- Determinar a causa de falhas, problemas e defeitos quase que em tempo real;
- Gerar cupons no ponto de venda com base nos hábitos de compra do cliente;
- Recalcular carteiras de riscos completas em minutos;
- Detectar comportamentos fraudulentos antes que eles afetem sua organização.

Com grandes quantidades de informações sendo transmitidas de inúmeras fontes, os bancos são desafiados a encontrar maneiras novas e inovadoras de gerenciar big data. Embora seja importante entender os clientes e aumentar sua satisfação, minimizar riscos e fraudes, além de manter uma conformidade regulatória, também é. O big data traz ótimos insights, mas exige que as instituições financeiras estejam um passo à frente com inteligência analítica avançada.

2.4. Algoritmos de Trading (AlgoTrading)

O trading por algoritmos, também chamado “trading automático” ou “trading de alta frequência”, utiliza um sistema de cálculo baseado em um modelo matemático: o algoritmo. Trata-se, portanto, de um software que determina as posições a serem tomadas, o sentido da tendência, o volume a ser negociado e, finalmente, a estratégia a ser adotada.

O trading por algoritmos apresenta, portanto, certas vantagens:

- Primeiramente, o trading por algoritmos permite aos investidores que não dispõem de tempo suficiente a dedicar à especulação nos mercados tomarem posições sem precisar acompanhar o mercado;

- Em segundo lugar, o trading automático permite uma presença nos mercados 7 dias por semana e 24 horas por dia, o que nenhum ser humano poderia fazer.
- Finalmente, o trading por algoritmos segue uma estratégia rigorosa, exclusivamente fundamentada na análise técnica dos mercados, e não sofre de nenhum estresse ou emoção passível de favorecer erros de avaliação.

O trading por algoritmos pode ser utilizado só ou em complemento a uma análise humana da situação de um mercado. Com efeito, utilizado sozinho, ele se constitui como um software de trading automático, já que é ele que posiciona as ordens no seu lugar, de acordo com os sinais detectados. Em contrapartida, utilizado em complemento a uma análise fundamentalista, ele constitui apenas uma ajuda à análise global e para a tomada de decisão, favorecendo uma estratégia específica.

Os parâmetros utilizados por esses algoritmos são muitos e difíceis de analisar detalhadamente. Evidentemente, eles levam em consideração a cotação atual de um ativo, mas também sua cotação histórica, a fim de determinar os diversos modelos de evolução possíveis.

Ele se baseia também nos volumes de vendas e compras de títulos utilizando assim dados da volatilidade para prever os momentos de tomada de posição mais oportunos. Além disso, o algoritmo integra no seu cálculo, os níveis de suporte e resistência técnica para gerar uma análise mais apurada do mercado.

O investimento automatizado só é possível graças a uma área chamada algorithmic trading, ou, algo trading, que está na intersecção de várias outras áreas, como computação, estatística, matemática, finanças, economia e aprendizado de máquina. O algo trading é o processo de realizar negociações por meio de um algoritmo, que nada mais é que uma receita de bolo ou uma sequência de instruções que o computador segue para atingir um certo objetivo, que neste caso é a execução automática de trades. Os algoritmos são interpretados e executados por computadores, que, de forma automática, recebem dados do mercado, os processam e geram sinais que disparam ordens para a Bolsa de Valores, tudo isso seguindo a definição feita pelo trader.

Uma automatização seguirá sempre as mesmas regras, independentemente de emoções, dificuldades que são recorrentes na rotina de um trader, às quais os robôs investidores são imunes.

É importante ter atenção especial em relação à infraestrutura do ambiente em que seus robôs investidores são executados. Se você utilizar uma plataforma de automatização de investimentos na nuvem, ou seja, que não depende de um programa instalado em seu computador ficar aberto e em funcionamento, mas que pode ser acessada online em qualquer lugar, terá maior confiança em relação à operação contínua do seu robô. Plataformas online para automatização de investimentos ficam hospedadas em servidores contratados para estarem funcionando a todo momento, garantindo alta disponibilidade, redundância e baixa latência, fazendo com que você não dependa do seu computador, da internet ou da energia elétrica de sua casa ou escritório.

2.5. Direct Market Access (DMA)

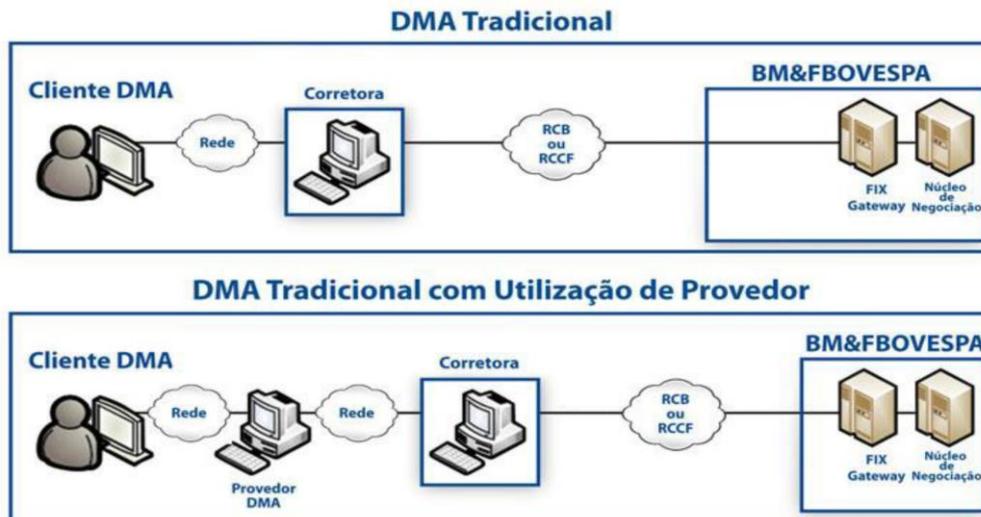
A BM&FBovespa possui vários modelos de negociação, um dos principais é o DMA (Direct Market Access), que permite ao cliente final o acesso direto ao mercado e ao ambiente da Bolsa no segmento de Equities (Bovespa) e Futuros (BM&F). O investidor, devidamente habilitado na BM&FBovespa e por intermédio de uma Corretora de Valores certificada e sob responsabilidade desta, pode negociar no sistema eletrônico da Bolsa, enviando ordens (oferta de compra e venda de ações) e recebendo em tempo real, as informações de difusão ao mercado, incluindo o livro de ofertas do sistema eletrônico de negociação.

Estão autorizados a operar neste modelo os Clientes Institucionais Individuais, Clientes Investidores Institucionais e Clientes Investidores Instituições Financeiras. Este modelo é composto por quatro tipos de conexões possíveis, são elas:

Modelo 1 – DMA Tradicional

Conexão que permite ao cliente negociar ações utilizando obrigatoriamente a infraestrutura tecnológica da Corretora para roteamento de ofertas para a BM&FBovespa no segmento BM&F e Bovespa.

Figura 1-1 – Modelo 1 – DMA Tradicional



Fonte: BM&FBovespa, 2011, p.32

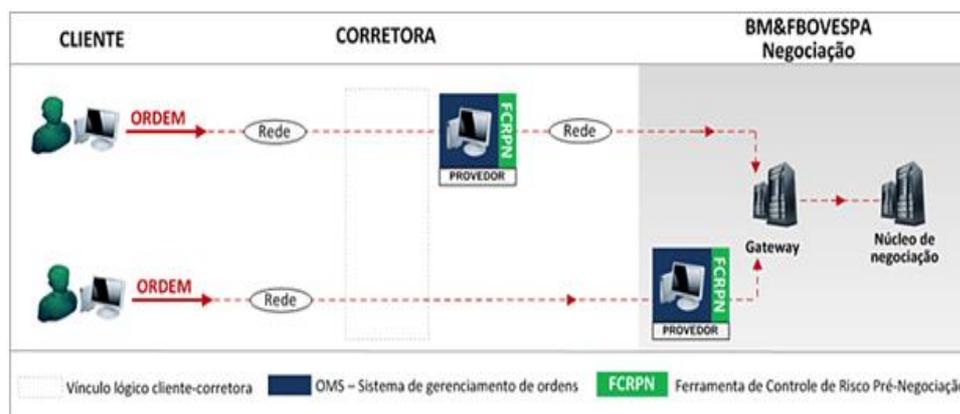
A primeira figura nos mostra que o cliente se conecta diretamente à infraestrutura tecnológica da corretora e, posteriormente, esta se conecta à BM&FBOVESPA. A segunda figura, por sua vez, ilustra que o cliente conecta-se à empresa prestadora de serviço de DMA, e esta empresa conecta-se à corretora, a qual, por fim, conecta-se à BM&FBOVESPA.

Esse modelo requer que a corretora possua, para controlar o acesso à infraestrutura tecnológica da BM&FBOVESPA de seus clientes às plataformas de negociação, sistema próprio de gerenciamento de ordens, também chamado de OMS (order management system).

Modelo 2 – DMA via provedor

Conexão que permite ao cliente negociar ações utilizando um provedor de DMA que se conecta diretamente à Bolsa. Esses provedores necessitam estarem certificados pela BM&FBovespa. O “vínculo lógico” entre o cliente e a corretora é controlado pelos sistemas da Bolsa e da empresa provedora de DMA – uma vez que as mensagens enviadas pelo cliente não trafegam pela infraestrutura tecnológica da corretora, mas sim pela infraestrutura da provedora de DMA.

Figura 1-2 – Modelo 2 – DMA via Provedor



Fonte: BM&FBOVESPA

Sungard Global Trading e Bloomberg Trade Book são alguns exemplos de fornecedores que proveem serviço neste modelo de negociação, no qual os clientes não se conectam a Corretora para negociar no mercado acionário;

DMA Modelo 3 – Via Conexão Direta

Neste modelo, os clientes enviam ofertas diretamente via conexão direta com a BM&FBovespa em nome de um participante do mercado habilitado na Bolsa. É importante ressaltar que, nesta modalidade não é utilizado à infraestrutura tecnológica da corretora ou de algum fornecedor habilitado pela Bolsa.

Figura 1-3 – Modelo 3 – DMA via Conexão Direta



Fonte: BM&FBOVESPA

Modelo 4 – DMA via Conexão Direta - Co-location

Segundo a BM&FBovespa, o ambiente de Co-location é uma área segregada em seu centro de processamento de dados (CPD), com acesso restrito e controlado e infraestruturas física e lógica para acesso aos sistemas de negociação por meio do uso de ATS, software responsável por realizar negociações no ambiente de Co- location.

Neste modelo, os clientes enviam ordens através de softwares instalados no data center da BM&FBovespa. O Investidor possui acesso remoto através de VPN nos equipamentos para realizar manutenção, configuração, monitoração e gerenciamento. A estrutura de servidores instalada no data center da Bolsa, pode ser do cliente ou da Corretora.

Segundo definição da BM&FBovespa, são oferecidas duas modalidades de Co- location para o mercado, a Co-Location Corretora e Co-Location Investidor.

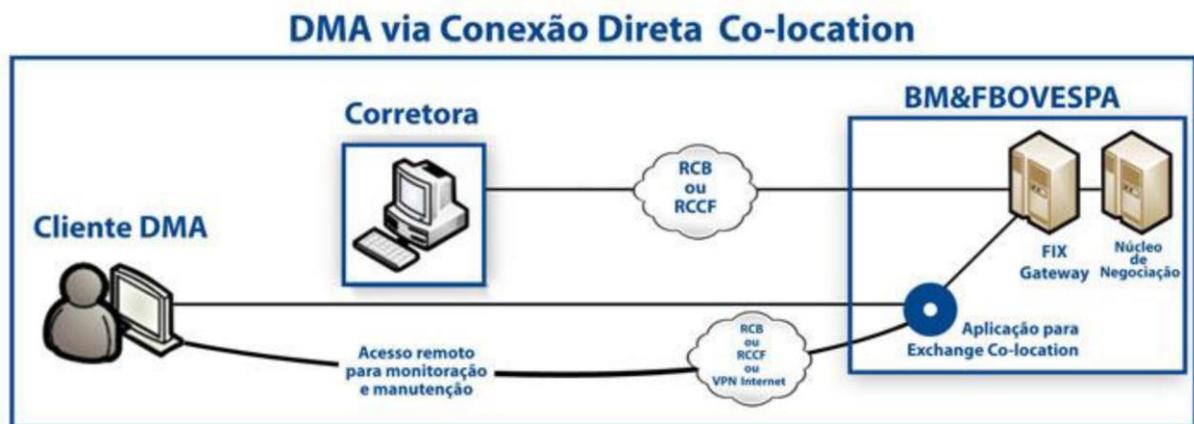
Co-location Corretora

- Modalidade de contratação do acesso pela corretora, cabendo somente a esta acessar a unidade de hospedagem e os equipamentos de negociação nela instalados para manuseio, administração, gestão e controle. A corretora poderá instalar em seus servidores softwares para envio de ordens tanto para carteira própria como para carteiras de clientes. Tais clientes, a critério da corretora, poderão acessar remotamente o(s) software(s) de negociação.
- As sessões de conectividade, estabelecidas para a negociação via Co-location, deverão destinar-se exclusivamente a esse fim. Caso sejam compartilhadas por diversos clientes, deverão ser distribuídas entre eles a critério da corretora.

Co-location Investidor

- Modalidade de contratação do acesso pelo investidor e sua corretora, cabendo somente ao primeiro acessar a unidade de hospedagem e os equipamentos de negociação nela instalados, para manuseio, administração, gestão e controle. O cliente poderá instalar em seus servidores softwares para envio de ordens exclusivamente em nome de carteiras sob sua gestão. Essa modalidade permite ao investidor negociar sob diversas corretoras utilizando um mesmo equipamento de negociação. Nesse caso, deverá ser estabelecida, no mínimo, uma sessão de conectividade para a negociação sob cada corretora, devendo cada sessão destinar-se exclusivamente a esse fim. Deve-se ressaltar que, independentemente da modalidade contratada, é vedada a sublocação de unidades de hospedagem.

Figura 1-5 – Modelo 4 – DMA via Conexão Direta



Fonte: BM&FBovespa, 2011, p.35

Um dos modelos de negociação que tem crescido exponencialmente nos últimos anos na BM&FBovespa é a utilização de estratégias de ações com robôs de execuções apoiados por algoritmos matemáticos construídos com foco em performance e baixa latência. Neste cenário e com a busca constante pelos investimentos em infraestrutura tecnológica, uma mão de obra cada vez mais qualificada nas Corretoras, e a disputa por novos investidores, garantem a expectativa de que os negócios através da alta frequência cresçam rapidamente.

Atualmente, a exigência dos clientes estrangeiros e locais em relação às Corretoras que proveem serviços de HFT (High-frequency trading) tem aumentando substancialmente. E a decisão por este ou aquele serviço é definida pelo tempo. A acirrada disputa por baixa latência acaba sendo o divisor de águas para que o investidor defina com qual parceiro negociará suas estratégias de ações e opções. Além deste ponto, a diversidade de estratégias, possibilidade de customização das mesmas, a alta disponibilidade de infraestrutura, a qualidade do serviço de pós-trading e a equipe de suporte ao cliente são as principais características analisadas pelos clientes na escolha da Corretora responsável pelo intermédio das operações.

Para auxiliar a propagação do mercado de alta frequência, a BM&F Bovespa reduziu o valor das taxas em emolumentos (custo cobrado pela Bolsa para negociação) baseados em volume negociado. Essa atitude visa atrair e proteger os investidores, pois os custos para operar HFT são altos e as margens de lucro destas operações são relativamente pequenas.

2.6. High Frequency Trading (HFT)

Negociação de alta frequência (NAF) ou em inglês: High-frequency trading (HFT) é uma forma primária de negociação algorítmica em finanças que surgiu em 1998. Especificamente, é o uso de ferramentas tecnológicas sofisticadas e algoritmos de computador para comercializar rapidamente commodities e ativos financeiros. As negociações de alta frequência usam propriedades estratégicas de negociação para realizar por meio dos computadores as negociações de venda e compra em escala num intervalo de menos de 11 segundos, chegando a escala de microsegundos.

Como resultado, a negociação de alta frequência tem um potencial de Sharpe (uma medida de risco e recompensa) dezenas de vezes maior do que as tradicionais estratégias de compra e venda. operadores de alta frequência competem tipicamente com outros que atuam neste meio de negociação e de forma diferente dos investidores de longo prazo.

Empresas que atuam em negociação de alta frequência

compensam as baixas margens com incríveis altos volumes de tradings, freqüentemente numeração na casa dos milhões. Tem sido argumentado que um incentivo – núcleo em grande parte do desenvolvimento tecnológico por trás da negociação de alta frequência é essencialmente front running, em que os prazos variados na propagação de ofertas são aproveitados por aqueles que têm acesso antecipado à informação.

Um corpo substancial de pesquisas argumenta que as negociações de alta frequência e comércio eletrônico colocam novos tipos de desafios para o sistema financeiro. Os algorítmicos e a negociação de alta frequência foram descobertos por terem contribuído à volatilidade no "Flash Crash" de 6 de maio de 2010, quando fornecedores de negociações de alta frequência rapidamente se retiraram do mercado. Vários países europeus propuseram cercear ou proibir as negociações de alta frequência devido a preocupações sobre a volatilidade. Outras queixas contra a negociação de alta frequência incluem o argumento de que algumas empresas deste meio raspam os lucros dos investidores quando os fundos de índice reequilibram as suas carteiras.

De acordo com uma pesquisa publicada no livro "A Finança Digitalizada: capitalismo financeiro e revolução informacional" cerca de 40% de todas as operações de compra e venda de ações na Bolsa de Valores brasileira em 2014 foram realizadas por programas de computador automatizados.

O autor da pesquisa e do livro é o sociólogo e pesquisador do Ipea (Instituto de Pesquisa Econômica Aplicada) Edemilson Paraná. Os dados revelam ainda que um "robô" é capaz de realizar uma operação a cada 10 milissegundos.

Com a expansão do mercado, várias corretoras e até bancos correm para oferecer esse tipo de serviço de negociação automatizada em algoritmos, ou seja, de alta frequência. Nos Estados Unidos, o HFT (High Frequency Trading) correspondeu a aproximadamente 70% de todo o volume negociado na bolsa durante 2014.

3. Métodos e procedimentos de coleta e de análise de dados do estudo

Este capítulo pretende informar sobre as diversas decisões acerca da forma como este estudo foi realizado.

3.1. Estratégia de investigação OU Métodos de pesquisa utilizados

Todos os dados usados para a realização deste estudo foram coletados no Bloomberg, no período de 02/01/2012 até 01/11/2018.

3.2. Procedimentos e instrumentos de coleta de dados utilizados no estudo

No Excel, configurado para extração de dados do sistema Bloomberg e através da seguinte fórmula:

```
=BDH(A4;"PX_Volume";"10/30/2004";"";"dir=h";"dts=o";"days=t";"cols=1731;rows=1")
```

Coletamos os dados referentes ao Volume diário das Ações que compõem a Carteira Téorica do Índice Ibovespa.

No Excel, configurado para extração de dados do sistema Bloomberg e através da seguinte fórmula:

```
=BDH(A2;"AVERAGE_BID_ASK_SPREAD";"10/30/2004";"";"dts=h";"dir=h";"days=t";"fill=0";"cols=1693;rows=1")
```

Coletamos os dados referentes à média diária do Bid-Ask spread das Ações que compoem a Carteira Téorica do Índice Ibovespa.

O Bloomberg já informa os dados em forma diária, devido a esse fato não foi necessário realizar nenhum cálculo para obter a média diária do Bid-Ask spread.

Com intuito de simplificação, selecionamos 5 ações com peso representativo sobre o Índice Ibovespa e de diferentes segmentos de mercado. São elas: PETR4 (Petróleo Brasileiro S.A.) ; ITUB4 (Itaú Unibanco) ; VALE3 (Vale) ; ABEV3 (Companhia de Bebidas das Américas S.A) e ELET6 (Eletrobras - Centrais Elétricas Brasileiras S.A.).

3.3. Formas de tratamento e análise dos dados coletados para o estudo

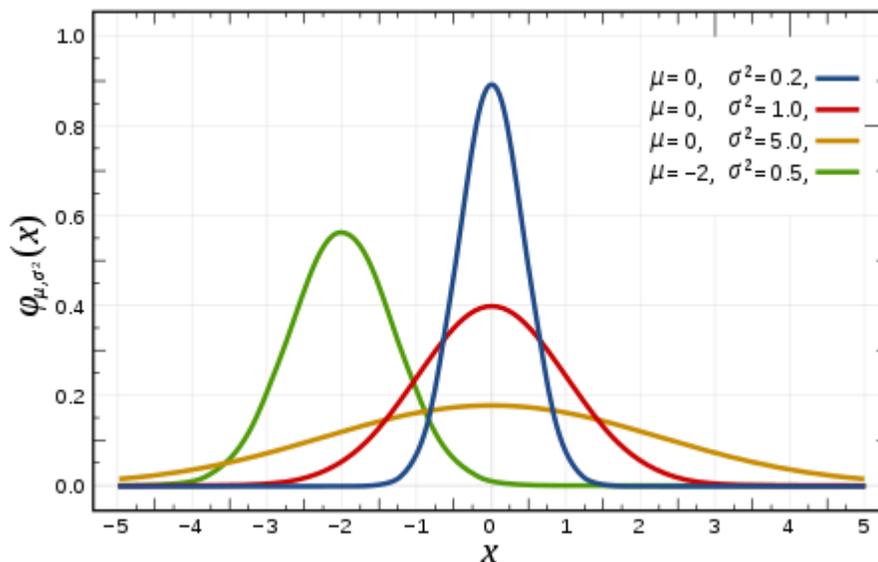
Após selecionarmos as respectivas ações calculamos o valor médio das 5 ações juntas, somando os valores e dividindo por 5 para se chegar à média do volume e média do Bid-Ask spread diário.

Dividiu-se a amostra em dois períodos exatamente iguais para realizar o teste estatístico Z para comparar as duas médias. Um teste Z é um teste de hipótese baseado na estatística Z, que segue a distribuição normal padrão sob a hipótese nula.

Teste Z é qualquer teste estatístico no qual a distribuição do teste estatístico sob a hipótese nula pode ser aproximada por uma distribuição normal. É um teste estatístico usado para inferência, capaz de determinar se a diferença entre a média da amostra e da população é grande o suficiente para ser significativa estatisticamente.

Por conta do teorema central do limite, muitos testes estatísticos são normalmente distribuídos para grandes amostras. Para cada nível de significância, o Teste Z tem um único valor crítico, o que torna o teste mais conveniente que o Teste T de Student que tem valores críticos separados para cada tamanho de amostra. Assim, muitos testes estatísticos podem ser realizados como testes Z aproximados se: o tamanho da amostra é grande ou a variância da população é conhecida. Se a variância da população é desconhecida (tendo que ser estimada a partir da amostra) e o tamanho da amostra é pequeno ($n < 30$), o teste T de Student pode ser mais apropriado.

Figura 2 - Função de densidade de probabilidade para a distribuição normal.



Fonte: Mathematica, Inkscape

A seguir foram feitos os cálculos de desvio padrão e variância da amostra dividida em dois períodos exatamente iguais.

Com o intuito de comparar as médias entre os dois períodos imputamos as informações no Teste Z.

3.4. Limitações do método

As informações da média do Bid-Ask spread só estão disponíveis no Bloomberg a partir de 02/01/2012, impossibilitando de realizar a comparação pretendida de 7 anos exatos por período = 2004 a 2011 e 2011 a 2018.

4. Apresentação e análise dos resultados

Este capítulo apresenta e discute os principais resultados alcançados, analisa e discute suas implicações e produz sugestões sobre o estudo previamente selecionado.

4.1 Análise do Average Bid-Ask spread

Começamos pela análise da comparação entre as duas médias (Teste Z) do Bid-Ask spread entre os dois períodos = 02/01/2012 a 03/06/2015 e 05/06/2015 a 01/11/2018. E como se era esperado, identificou-se que a média do primeiro período é superior à média do segundo e podemos afirmar isso estatisticamente uma vez que o valor de $Z = 1,969771548$ é superior ao z crítico uni-caudal = $1,644853627$, confirmando a hipótese.

Tabela 1.1 – Teste Z para o Average Bid-Ask spread

Teste-z: duas amostras para médias

| | <i>Variável 1</i> | <i>Variável 2</i> |
|--------------------------------|-------------------|-------------------|
| Média | 0,024146879 | 0,021024066 |
| Variância conhecida | 0,002114 | 0,00001233 |
| Observações | 846 | 846 |
| Hipótese da diferença de média | 0 | |
| z | 1,969771548 | |
| $P(Z \leq z)$ uni-caudal | 0,024432279 | |
| z crítico uni-caudal | 1,644853627 | |
| $P(Z \leq z)$ bi-caudal | 0,048864559 | |
| z crítico bi-caudal | 1,959963985 | |

Sendo a Variável 1 correspondente ao período de 02/01/2012 a 03/06/2015 e a Variável 2 correspondente ao período de 05/06/2015 a 01/11/2018.

4.2 Análise da evolução do Volume

A segunda análise, já no âmbito do Volume diário, também testa a comparação entre as duas médias (Teste Z) nos mesmos períodos, 02/01/2012 a 03/06/2015 e 05/06/2015 a 01/11/2018. E mais uma vez como já era esperado observamos que a média do primeiro período é superior à média do segundo e podemos afirmar isso estatisticamente uma vez que o valor de $Z = 20,06478456$ é superior ao z crítico uni-caudal = 1,644853627, confirmando a hipótese.

Tabela 1.2 – Teste Z para a evolução do Volume

Teste-z: duas amostras para médias

| | <i>Variável 1</i> | <i>Variável 2</i> |
|--------------------------------|-------------------|-------------------|
| Média | 18596753,59 | 12031846,93 |
| Variância conhecida | 3,6966E+13 | 5,35986E+13 |
| Observações | 846 | 846 |
| Hipótese da diferença de média | 0 | |
| z | 20,06478456 | |
| $P(Z \leq z)$ uni-caudal | 0 | |
| z crítico uni-caudal | 1,644853627 | |
| $P(Z \leq z)$ bi-caudal | 0 | |
| z crítico bi-caudal | 1,959963985 | |

Sendo a Variável 1 correspondente ao período de 05/06/2015 a 01/11/2018 e a Variável 2 correspondente ao período de 02/01/2012 a 03/06/2015. Repara-se que invertemos os períodos para realizar a análise do Volume.

5. Conclusões e recomendações para novos estudos

Conclui-se com o estudo em questão que de fato se compararmos dois períodos iguais 02/01/2012 a 03/06/2015 e 05/06/2015 a 01/11/2018 tivemos um aumento significativo da média diária dos volumes das 5 ações selecionadas na amostra que é suficiente para ser significativa estatisticamente, e também uma diminuição significativa do Bid-Ask spread diário também das 5 ações selecionadas na amostra, o que pode indicar uma eficiência em termos de liquidez para o Mercado de Capitais Brasileiro.

No entanto, não podemos afirmar que a introdução da Inteligência Artificial, Machine learning, Big Data, Algotradings, Co-location e HFT contribuíram de maneira isolada e independente para a liquidez do mercado de capitais Brasileiro. Outros fatores macro- e micro-econômicos também podem ter contribuído para não só a evolução do volume mas também para o encurtamento do Bid-Ask spread.

Vale ressaltar que não foi utilizado o período anterior a 2012 na análise por uma questão de limitação do acesso aos dados, mas como sugestão para novos estudos seria indicado realizar o mesmo teste estatístico para os períodos de 2004 a 2011 e 2011 a 2018, exatos 7 anos da implementação do DMA 3 e 4 pela hoje conhecida [B]³.

6. Referências Bibliográficas

Estratégia-Bolsa. **AlgoTrading: O que é o trading por algoritmos?** Disponível em: <<https://www.estrategia-bolsa.pt/trading-por-algoritmos.html>>. Acesso em 2 de Julho de 2018.

HAYDN, S. **Gestores de Riscos estão preparados para a Inteligencia artificial?** GARP 2017. Disponível em: <<http://www.garp.org/#!/riskintelligence/all/all/a1Z40000003PIkaEAG>> Acesso em Fevereiro de 2017.

LANA, MATEUS. SmarttBot. **Investimento automatizado na Bolsa: tudo o que você precisa saber.** Disponível em: <<https://smarttbot.com/blog/investimento-automatizado-na-bolsa-o-que-voce-precisa-saber/>>. Acesso em 20 de Fevereiro de 2016.

MICHAEL, L. **Flash boys : A Wall Street Revolt.** New York: W. W. Norton & Company, 2014.

Minitab. **Teste Z: O que é um teste Z?** Disponível em: <<https://support.minitab.com/pt-br/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/tests-of-means/what-is-a-z-test/>>. Acesso em 20 de Novembro de 2018.

SAS. **Big Data: O que é e qual sua importância?** Disponível em: <https://www.sas.com/pt_br/insights/big-data/what-is-big-data.html>. Acesso em 8 de Fevereiro de 2017.

SAS. **Inteligencia Artificial: O que é e qual sua importância?** Disponível em: <https://www.sas.com/pt_br/insights/analytics/inteligencia-artificial.html>. Acesso em 10 de Maio de 2017.

SAS. **Machine Learning: O que é e qual sua importância?** Disponível em: <https://www.sas.com/pt_br/insights/analytics/machine-learning.html#machine-learning-importance>. Acesso em 15 de Junho de 2017.

SÉRGIO, A. P. **Como atender clientes HFT (HighFrequency Trading) nas operações de estratégias no mercado de ações com agilidade e eficiência?.** Rio de Janeiro: FGV, 2013.

SOUZA, PABLO. **Conheça os modelos de acesso direto ao mercado – BM&FBOVESPA: DMA 1, 2, 3 e 4** Disponível em: <<https://blog.cedrotech.com/conheca-os-modelos-de-acesso-direto-ao-mercado-bmfbovespa-dma-1-2-3-e-4/>>. Acesso em 20 de Agosto de 2016.