

Avaliações de tendência de ações da bolsa de valores através do algoritmo de regressão linear

Stock Market trend assessments using the Linear Regression algorithm

Lucas Vinicius de Oliveira
Centro Universitário Hermínio Ometto – UNIARARAS
Estudante de Engenharia da Computação do Centro Universitário Hermínio
Ometto – UNIARARAS
Lucas_araras@hotmail.com

RESUMO

Hoje em dia várias pessoas sonham com a independência financeira e um dos meios de se alcançar é com investimento em ações na bolsa de valores, um mercado que “enche os olhos” a primeira vista mas, é preciso muita cautela ao se aventurar nesse meio. É exigido um nível de conhecimento muito alto para se operar nesse tipo de mercado, saber o que comprar, vender e em quais momentos, são cruciais para um resultado positivo. O objetivo deste artigo é propor um motor de investimentos que ajude na tomada de decisões na hora da compra ou venda, analisar dados históricos da bmf Bovespa e fazer previsões calculadas com base teórica para se obter o sucesso na operação e aumentar os lucros. Será abordado todas as etapas de estudos feitos para o resultado proposto e algoritmo utilizado, bem como o conceito

de data mining por trás deste trabalho no auxílio da ferramenta de um motor de investimentos.

Palavras-chave: *Linear Regression, Bolsa de valores, Ações, Mercado financeiro*

INTRODUÇÃO

Acredita-se que a bolsa de valores tenha surgido em meados de 1487, em Bruges, na Bélgica, mas, as primeiras ações que se tem notícia foram em 1602 na bolsa de Amsterdã, no Brasil o primeiro histórico que se tem foi em 1845 na bolsa que se instalou na cidade do Rio de Janeiro.

O maior problema nos dias atuais é a volatilidade do mercado que fazem inúmeras pessoas desistirem por terem prejuízos nas suas operações. Atualmente existem diversas ferramentas de investimento automatizados por robôs como por exemplo o SmarttBot

Figura 1 – site do SmarttBot (robô de investimentos).



Neste artigo é apresentado uma solução usando o algoritmo de Regressão Linear. A ideia principal é fazer um estudo com dados históricos coletados através do site da infomoney do mês de Abril de 2019.

Filtramos para estudo o ativo mini índice que consiste em um contrato futuro do índice Bovespa. Segundo o site do Toro Radar:

“Ele nada mais é do que uma carteira teórica de ações, composta pelas maiores empresas que compõe a bolsa em um só contrato. Ou seja, ele agrupa as valorizações e desvalorizações dos papéis que mais movimentam dinheiro na bolsa. Com isso, ele se torna um índice que é utilizado para se comparar todos os outros investimentos do mercado de renda variável. Em termos técnicos dizemos que esta carteira teórica é um benchmark para o mercado.”
(<https://www.tororadar.com.br/>)

Com os dados em mãos e usando a ferramenta Weka, iremos prever tendencias deste ativo e a idéia final é mostrar ao investidor se no dia subsequente será melhor a opção de compra ou Venda. Nas próximas etapas do artigo será abordado quais dados existem no arquivo, o que significa cada um e sua influencia nos estudos e previsões futuras.

REFERENCIAL TEÓRICO

Data Mining:

Com a ideia de se extrair informações importantes em um grande volume de dados foi criado o data mining. Segundo GOLDSCHMIDT e PASSOS (2015), embora não exista um consenso quanto a esta diversidade terminológica, todas se referem

a tecnologia e aplicações cujo objetivo é identificar um conhecimento útil a partir de grandes volumes de dados.

Esse tipo de aplicação foi bem útil e ajudou inúmeras empresas a organizarem os seus dados, obtendo padrões relevantes para futuras melhorias. A NASA por exemplo utilizou deste método por conta dos trilhões de dados gerenciados, a IoT (*internet of things*) é outro exemplo bastante comum.

De acordo com GOLDSCHMIDT e PASSOS (2015):

“Atualmente, dados científicos em projetos de pesquisa, tais como missões espaciais da NASA e o Projeto do Genoma Humano, tem alcançado proporções gigantescas. Alguns estudos estimam uma taxa de crescimento de dados mundial em torno de 40% ao ano, a próxima década, o que, em 2020, deverá alcançar um total de cerca de 44 zettabytes de informações digitais em todo o mundo. O Advento da IoT (sigla para a expressão em ingles Internet of Things), por exemplo, tem certamente uma participação expressiva nesse crescimento.”

Contudo, podemos observar o poder desta ferramenta em um futuro muito próximo, onde tudo tende-se a se conectar facilitando a vida pessoal ou da indústria.

História dos dados:

Como comentado anteriormente, neste trabalho utilizou-se uma base de dados do mês de Abril/2019 do ativo mini-índice, neste arquivo dispomos de seis variáveis, sendo elas:

- Data
- Ultimo valor (em pontos)
- Abertura (em pontos)
- Máxima (em pontos)
- Mínima (em pontos)

- Saída (binário)

Data: Data do ativo negociado, somente dias da semana, finais de semana e feriados não são operáveis.

Ultimo Valor: É o valor em pontos que o ativo fechou no dia anterior.

Abertura: Valor em pontos em que o ativo começou o dia corrente.

Máxima: Valor máximo em pontos que o ativo atingiu no dia corrente.

Mínima: Valor mínimo em pontos que o ativo atingiu no dia corrente.

Saída: Consiste no valor binário (zero ou um) para a indicação, se foi um dia que a abertura foi maior que o fechamento temos zero que se refere a venda, o contrário é o número um que se refere a compra.

Figura 2 – Dados do ativo mini indice extraídos em arquivo .arff no weka.

Data	Fechamento Dia anterior	Abertura do dia	Máxima do dia	Mínima do dia	Saída (binário)
30-04-2019	9702500	9710500	9742000	9624000	0
29-04-2019	9693200	9709500	9784000	9667500	0
26-04-2019	9689300	9689000	9746500	9626000	1
25-04-2019	9716400	9599500	9726000	9483000	1
24-04-2019	9570000	9710000	9721000	9481000	0
23-04-2019	9671600	9570000	9712500	9565000	1
22-04-2019	9534700	9500000	9582000	9441000	1
18-04-2019	9522300	9490000	9605000	9414000	1

MATERIAIS E MÉTODOS

Weka (*Waikato Environment for Knowledge Analysis*), teve início em 1993 desenvolvido em uma falcudade da Nova Zelândia chamada Waikato.

O software tem o intuito de agregar vários algoritmos com ligações na área de inteligencia artificial e aprendizagem de máquina (*machine learning*), técnica essa que representa a capacidade do computador “aprender” com uma determinada massa de dados.

Trabalha através de padrões tentando encontrar a maneira mais correta e que mais represente a teoria correta para seguir em um fluxo de padrões e repetições com os dados apresentados.

Neste trabalho utilizaremos o weka como ferramenta principal no auxílio da aprendizagem dos dados obtidos do ativo mini-índice, juntamente com o algoritmo de Regressão Linear para reconhecimento de padrões e, assim a chance de aprender com os dados coletados para possíveis privisões do ativo.

Figura 3 – Tela inicial do software weka.



O trecho abaixo explica o algoritmo de regressão linear de uma forma bem intuitiva.

Segundo Marcelo Petenati (2015):

“Eu posso não ter muitas informações sobre você, mas se me disser sua altura, posso ter uma ideia do seu peso. Se me disser sua idade, posso lhe indicar uma música que provavelmente gostará. Se me disser sua escolaridade, posso ter uma estimativa de quanto você ganha por mês.

Pode parecer mágica, mas isso é apenas uma amostra do que é regressão linear. O método nada mais é do que estudar, em números, a relação entre variáveis. Você, com certeza, já visitou um site de compras e, ao clicar em um item, logo abaixo, encontrou a opção “Você provavelmente vai gostar desses outros produtos”.

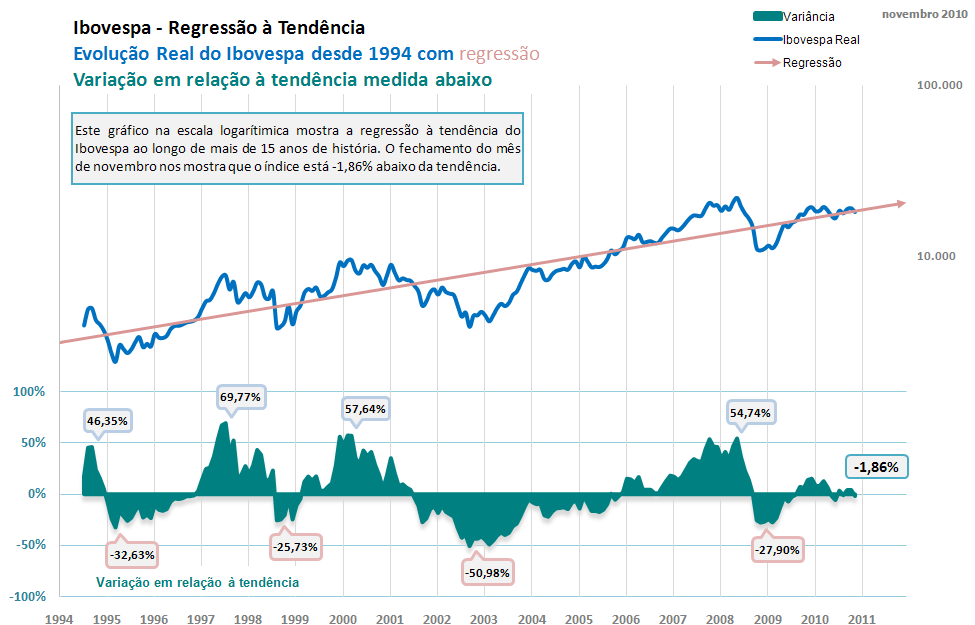
Outro exemplo importante: se me disser qual foi o último filme que assistiu, posso lhe indicar um que com certeza gostará. É assim que funciona quando você vê um filme ou uma série na Netflix e, logo depois, ela lhe indica um outro título que surpreendentemente lhe agrada.

Isso são algoritmos que conseguem estimar o que você pode vir a gostar com base no que você já gostou ou visitou. Essa é a ideia central de uma regressão linear. Ou seja, estudar o que temos para poder entender como reagir no futuro. O método garante um aspecto importante no mercado hoje: a entrega de produtos cada vez mais customizados.”

Com o mercado de ações não é diferente, pois o mercado sempre relembra seu passado, é como se ele “gostasse” dos pontos marcantes que já chegou e assim um dia voltará neles. Uma ótima forma de operar suporte e resistência de mercado, por exemplo.

Abaixo uma foto de regressão a tendência do ativo do Ibovespa e sua variação:

Figura 4 – Gráfico de regressão a tendência do ativo Ibovespa.



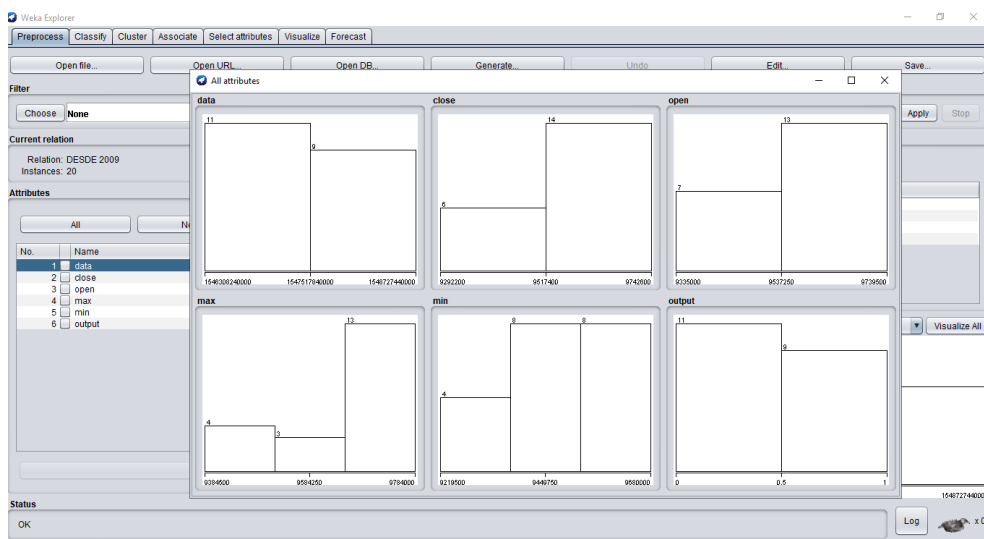
Segundo Marcelo Petenati (2015), Depois de ter uma ideia do que é o conceito por trás do método, trazendo para um cenário mais formal, fazer um modelo de regressão linear é estudar uma variável de interesse (variável dependente) em função de outras variáveis que me ajudarão a entendê-la — as chamadas covariáveis ou variáveis independentes.

RESULTADOS

Foi oferecido como base de dados para o weka o range de dias de 01/04/2019 a 22/04/2019 deixando os dias 23,24,25,26,29 e 30/04/2019 para ser os dias de teste da análise do algoritmo de regressão linear.

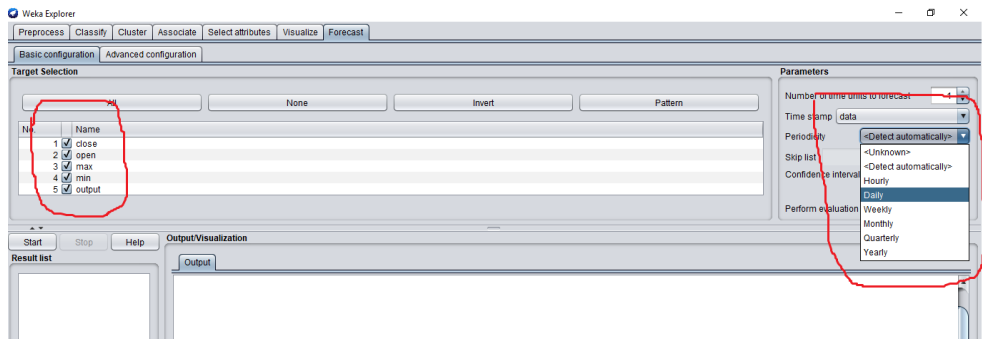
Para a base de dados ser reconhecida na ferramenta foi modificado o arquivo original com extensão .xls para .arff, o resultado é exibido abaixo:

Figura 5 – Tela do Weka com os atributos do arquivo de dados separado por quantidades.



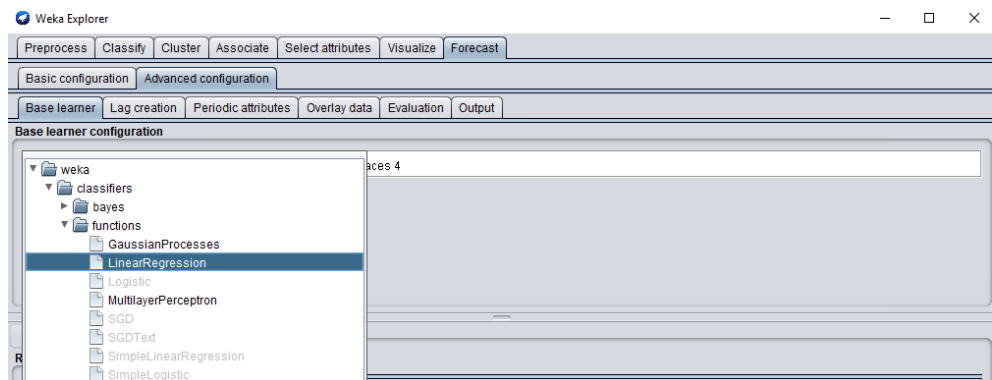
Após o tratamento do arquivo houve a necessidade do download da ferramenta “Forecast” para a predição correta dos dados. Essa ferramenta não é padrão do Weka e pode ser encontrada na guia Tools – Package Manager para instalação, é nela que encontraremos também o algoritmo de regressão linear.

Figura 6 – Ferramenta Forecast do software Weka.



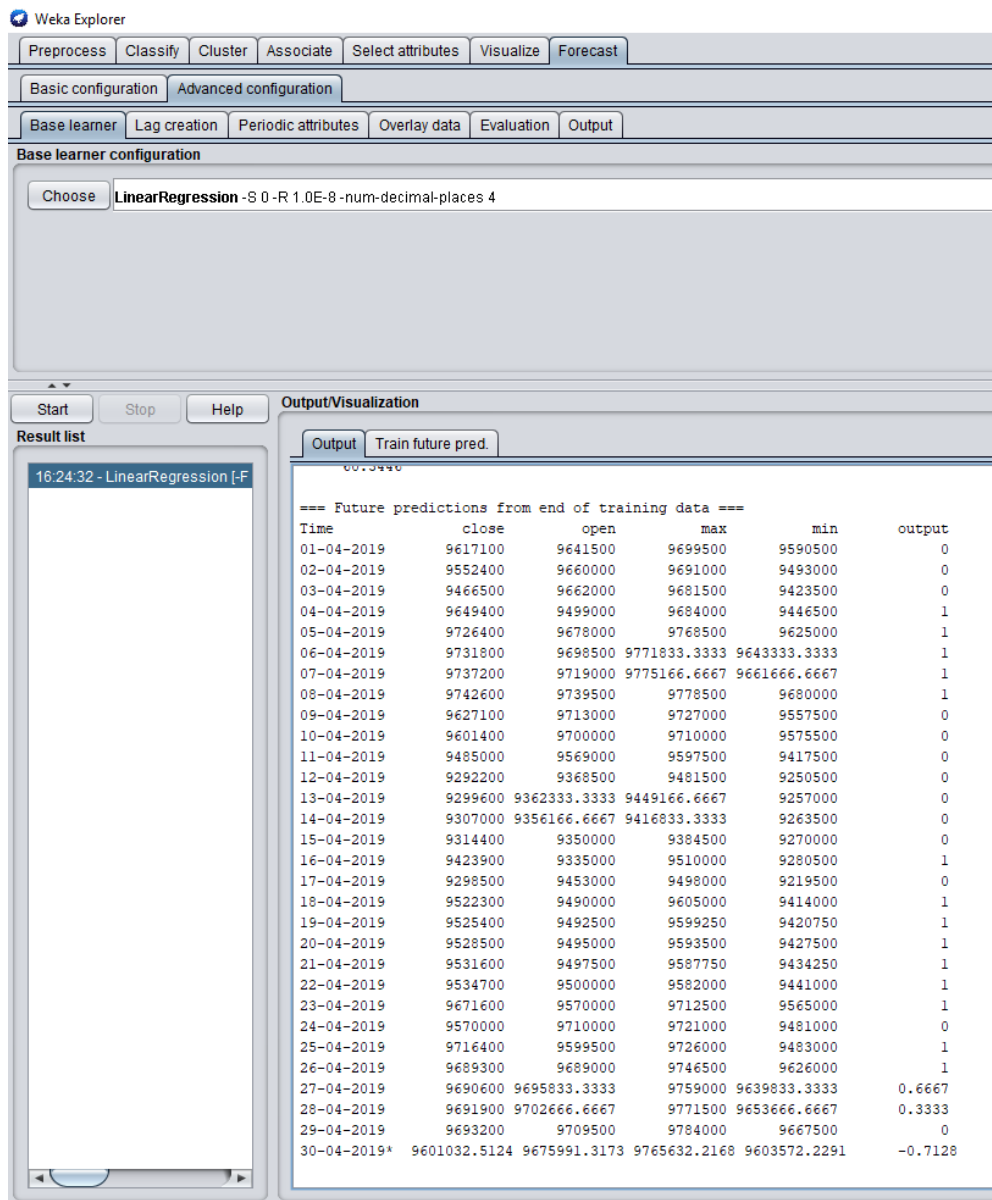
Na tela de Forecast é possível selecionar o Time Stamp que seria basicamente o “por onde” vamos começar e “qual variável” vamos deixar evidente na hora da geração do algoritmo. Nesse caso foi escolhido a data que afinal, é o nosso marcador de tempo, “no dia seguinte eu compro ou vendo?” Essa é a pergunta que nos fazemos para intuito deste trabalho.

Figura 7 – Tela de escolha do algoritmo a ser utilizado na aba Forecast do Weka.



Abaixo um exemplo do dia 30/04/2019 onde pelo algoritmo de regressão linear apresentou o resultado tendendo a venda do ativo.

Figura 8 – Demonstração de execução do algoritmo de regressão linear para o dia 30/04/2019 do arquivo de dados.



Se analisarmos o gráfico disponibilizado pelo site da investing.com vemos claramente a tendência de baixa do ativo nesse dia:

Figura 9 – Gráfico real do ativo mini índice no dia 30/04/2019 extraído do site investing.com



Ainda que o algoritmo não vise a melhor saída no ativo teríamos tido um ganho de 105 pontos aproximadamente, convertidos em reais seriam algo em torno de R\$21,00

Após os testes feitos obtivemos o resultado como é mostrado na tabela abaixo:

Figura 10 – Resultados finais obtidos após rodar o algoritmo de regressão linear para os dias 22,23,24,25,26,29 e 30/04/2019.

RESULTADOS OBTIDOS				
DIA	RESULTADO			
22-04-2019,9534700,9500000,9582000,9441000,1	OK			
23-04-2019,9671600,9570000,9712500,9565000,1	OK			
24-04-2019,9570000,9710000,9721000,9481000,0	FALHOU			
25-04-2019,9716400,9599500,9726000,9483000,1	OK			
26-04-2019,9689300,9689000,9746500,9626000,1	FALHOU			
29-04-2019,9693200,9709500,9784000,9667500,0	FALHOU			
30-04-2019,9702500,9710500,9742000,9624000,0	OK			

57,14%

Analisando a imagem 10 podemos observar que tivemos 57,14% de acertos em nossos testes, um valor razoavelmente baixo, mas que para geração de lucro, atendeu as expectativas propostas.

Se fizermos o comparativo em reais, teríamos o seguinte cenário:

- 1 mini contrato = R\$0,20 * numero de pontos
- 100 pontos = R\$20,00 com 1 contrato

Podendo chegar a um lucro mensal de R\$ 1.114,00 com 100 pontos no dia e um único contrato do ativo mini indice.

Descontando os impostos pagos, que se chamam corretagem (valor pago já no dia seguinte as operações de day trade), demais impostos e 15% de imposto de renda sobre o lucro no final do mês em que se está operando o ativo.

Mas para isso precisa-se de muito estudo e didicação, horas na frente da tela, conhecer as tendências. Ou conhecer o nosso

exemplo de motor de investimento usando regressão linear.

CONCLUSÃO

É possível perceber que em uma ferramenta ou motor de investimentos criados, não vão ter cem por cento de sucesso em suas operações. O mercado é sempre imprevisível por mais que respeite padrões e indicadores relevantes, em épocas de eleições presidenciais por exemplo, isso fica ainda mais evidente.

A solução proposta respondeu positivamente ficando com porcentagem acima dos cinquenta por cento para que fosse gerado um lucro, nem que sendo mínimo para o investidor que venha a utilizar esta ferramenta.

Os Objetivos futuros para este trabalho visam melhorar a performance do algoritmo, mudando parametros internos da ferramenta weka por exemplo para que se alcance uma maior porcentagem no resultado. Além disso, aprimorar as variáveis para que seja possível uma melhor saída do ativo no decorrer do dia, como já foi observado existem ganhos, mas não os melhores ganhos que poderiam ter sido concluídos se saímos da posição do ativo algumas horas antes do final do dia.

BIBLIOGRAFIA

K.K. Sureshkumar & Dr.N.M.Elango: **“An Efficient Approach to Forecast Indian Stock Market Price and their Performance Analysis”**. International Journal of Computer Applications (0975 – 8887) Volume 34– No.5, November 2011

Ronaldo GOLDSCHMIDT, & Emmanuel PASSOS: **“Data Mining: um guia prático”**. Editora Campus, Rio de Janeiro: Elsevier, 2005

CAMILO, C. O. & SILVA, J.C **“Mineração de Dados: Conceitos, Tarefas, Métodos e Ferramentas”**. 2009. Instituto de informática, Universidade Federal de Goiás, 2009

BASTOS, A. P. Z. et al. **“Utilização de um Jogo Sérió e Naïve Bayes para Auxiliar na Avaliação Cognitiva do Transtorno de Déficit de Atenção/Hiperatividade”**. 2012. Núcleo de Tecnologia Educacional de Niterói, 2012

MATTOS, P., 1998. **“Uso de testes computadorizados em neuropsicologia”**.
In:Gagliardi,R.J.; Reimao, R. (org.). Clínica neurológica. São Paulo, Lemos
Editorial, pp. 319-322.

O que é Regressão Linear? Entenda aqui!. Disponível em:
< <https://www.escolaedti.com.br/o-que-e-regressao-linear-entenda-aqui/>>. Acesso em: 09 abr. 2019.

História da Bolsa de Valores. Disponível em:
<http://www.mercadocomum.com/site/artigo/detalhar/historia_da_bolsa_de_valores/materias-publicadas>. Acesso em: 12 fev. 2019.

Minicontrato Futuro de Ibovespa na BM&F. Disponível em:
< <https://br.advfn.com/investimentos/futuros/mini-ibovespa>>. Acesso em: 12 fev. 2019.

Padrões Gráficos – Figuras de Continuidade e Reversão. Disponível em: < <http://daltonvieira.com/padroles-graficos-figuras-de-continuidade-e-reversao>>. Acesso em: 03 fev. 2019.

Conceitos e Técnicas sobre Data Mining. Disponível em: < <https://www.devmedia.com.br/conceitos-e-tecnicas-sobre-data-mining/19342>>. Acesso em: 10 nov. 2018.

Mineração de dados com WEKA, Parte 1. Disponível em: < <https://www.ibm.com/developerworks/br/opensource/library/os-weka1/index.html>>. Acesso em: 10 nov. 2018.