**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS  
NÚCLEO DE EDUCAÇÃO A DISTÂNCIA**

**Pós-graduação *Lato Sensu* em Ciência de Dados e Big Data**

**Rodrigo Schinoff Oliveira**

**COMPORTAMENTO DO MERCADO NA ABERTURA DO PREGÃO**

Belo Horizonte

2020

**Rodrigo Schinoff Oliveira**

**COMPORTAMENTO DO MERCADO NA ABERTURA DO PREGÃO**

Trabalho de Conclusão de Curso apresentado ao Curso de Especialização em Ciência de Dados e Big Data como requisito parcial à obtenção do título de especialista.

Belo Horizonte

2020

**SUMÁRIO**

1. Introdução 4

1.1. Contextualização 4

1.2. O problema proposto 5

2. Coleta de Dados 6

3. Processamento/Tratamento de Dados 7

4. Análise e Exploração dos Dados 8

5. Criação de Modelos de Machine Learning 8

6. Apresentação dos Resultados 8

7. Links 9

REFERÊNCIAS 10

1. Introdução

1.1. Contextualização

Nessa seção você deverá introduzir e contextualizar o seu problema de interesse.

No mercado financeiro, não podemos afirmar e muito menos garantir qual será a direção do preço de um determinado ativo no seu próximo movimento. No entanto, é possível estimar a direção do preço baseando-se na leitura do histórico dos movimentos gerados e principalmente verificando a tendencia do mercado.

Segundo a teoria formulada por Charles Henry Dow, os mercados movem-se em tendências, podendo ser de alta ou de baixa. As tendências por sua vez, podem ser primárias, secundárias e terciárias. Sendo a principal delas a tendência primária, pois representa o movimento mais longo do mercado. As tendências secundárias e terciárias correspondem as correções que ocorrem dentro da primária, ou seja, são pequenas interrupções ou desvios da tendência inicial, podendo retomar ao movimento anterior e dar seguimento a tendência de alta ou de baixa.

Se analisarmos em nosso cotidiano, podemos observar em vários aspectos formações de tendências. Durante as estações do ano, as temperaturas tendem a oscilar dentro de uma faixa esperada. No inverno por exemplo, as temperaturas provavelmente se manterão mais baixas do que na primavera ou verão, mas é muito possível que faça calor em um ou dois dias, ou até mesmo em uma semana inteira durante o inverno, mas que após este período as temperaturas retornem a média esperada para aquela estação. No trânsito, podemos observar um comportamento semelhante. Suponhamos que um determinado indivíduo percorra o mesmo trecho todos os dias no mesmo horário, em média ele deve levar um tempo muito próximo em seu deslocamento diário, porém alguns fatores como acidente de transito, via interrompida, aumento de fluxo devido a proximidade de um feriado prolongado, etc, certamente acarretarão no aumento do tempo para percorrer aquela mesma distância. O contrário também é verdadeiro, nos dias em que por alguma razão o trânsito apresentar-se mais livre, o tempo gasto no percurso será menor.

Em ambos os casos, as distorções que acarretam um desvio da média são ocasionadas por eventos isolados ou de pouca duração, porém os valores tendem a retornar para a sua faixa média. A probabilidade de uma retomada do comportamento médio inicial é bem maior do que a continuação ou o aumento de tais distorções.

1.2. O problema proposto

Dentro de uma tendência de mercado, o preço tende a acompanhar um movimento médio. No mercado financeiro, um dos indicadores utilizados na análise técnica para o acompanhamento médio dos preços de um determinado ativo, são as médias móveis. Ao longo de uma tendência, distorções podem ocorrer, fazendo com que o preço de um ativo se afaste dos valores médios. O movimento mais provável é, a diminuição desta distorção do que o aumento da mesma.

O objetivo deste trabalho é, analisar padrões nos movimentos de correções existentes imediatamente após a abertura do pregão, principalmente nos casos onde os preços iniciais definidos durante o leilão de abertura apresentam-se afastados de suas médias móveis.

A base analisada consiste na cotação do valor de fechamento do preço ocorrido no intervalo de cinco minutos da ação PETR4 da empresa Petróleo Brasileiro S.A – Petrobras no período de 09/09/2019 a 19/02/2020 e suas respectivas médias móveis de 17 e 72 períodos.

Nessa seção você deverá descrever o problema que você está propondo resolver através dos dados. Uma dica interessante é utilizar a técnica dos [5-Ws](https://its.unl.edu/bestpractices/remember-5-ws). Responder algumas dessas perguntas te dará uma melhor visão do seu problema e da solução.

(Why?) Por que esse problema é importante?

(Who?) De quem são os dados analisados? De um governo? Um ministério ou secretaria? Dados de clientes?

(What?): Quais os objetivos com essa análise? O que iremos analisar?

(Where?): Trata dos aspectos geográficos e logísticos de sua análise.

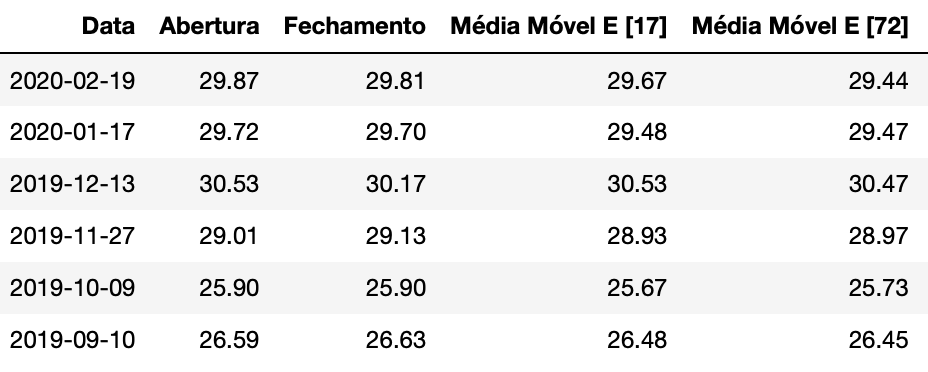
(When?): Qual o período está sendo analisado? A última semana? Os últimos 6 meses? O ano passado?

2. Coleta de Dados

Os dados desta análise foram extraídos de uma ferramenta privada de mercado, que possibilita a extração de cotações do mercado financeiro. Porém, as mesmas informações também podem ser adquiridas gratuitamente através do link  
<https://br.financas.yahoo.com/quote/PETR4.SA/history?p=PETR4.SA>

Nessa seção você deve deixar claro onde obteve os dados, o formato e estrutura dos datasets, o relacionamento entre os datasets utilizados, etc. Caso os dados tenham sido obtidos na internet, informe a data e o link em que os dados foram obtidos. Sugere-se que você crie uma tabela com a descrição de cada campo/coluna do seu dataset conforme o exemplo a seguir:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data e Hora | Preço Abertura | Preço Fechamento | Média Móvel 17 Períodos | Média Móvel 72 Períodos |
| 10/09/2019 10:05 | 26,59 | 26,63 | 26,48 | 26,45 |
| 07/10/2019 10:05 | 25,92 |  | 25,99 | 25,99 |
| 27/11/2019 10:00 | 29,01 |  | 28,93 | 28,97 |
| 16/12/2019 10:10 | 29,70 |  | 29,65 | 29,77 |
| 17/01/2020 10:10 | 29,72 |  | 29,48 | 29,47 |
| 19/02/2020 10:05 | 29,87 |  | 29,67 | 29,44 |



3. Processamento/Tratamento de Dados

Nessa seção você deve deixar registrado todo o processamento e tratamento feitos sobre os dados obtidos. É importante que você informe a quantidade de registros obtidos, a quantidade de registros duplicados ou com informações ausentes, que tratamento você deu para cada problema encontrado em seus datasets, etc. Você deve descrever cada passo de forma minuciosa, de forma que outra pessoa consiga reproduzir o seu processamento/tratamento de forma precisa. Justifique as decisões tomadas no tratamento dos dados. Por exemplo: para os valores ausentes para o campo X eu decidi preenchê-los utilizando o cálculo da média aritmética pelo motivo ..., e então justifique sua decisão.

A base de dados foi formada originalmente por registros diários gerados a cada cinco minutos, durante o intervalo do pregão do mercado a vista. O horário de funcionamento do mercado a vista é das 10:00hs às 16:50hs, gerando um total de 10043 registros. Estes registros são compostos de Data e Hora, Preço de Abertura, Preço de Fechamento, Preço Máximo, Preço Mínimo, Preço de Fechamento, Média Móvel Exponencial de 17 Períodos e Média Móvel Exponencial de 72 Períodos.

Foram acrescentadas duas novas colunas AbriuLongeMedia e RetornouMedia, com a finalidade de identificar os registros que tenham iniciado o empregão com preços distantes de suas médias móveis e que minutos depois tenham se aproximado das mesmas. São considerados preços longe das médias, os registros avaliados na abertura do mercado, ou seja, registros gerados até as 10:15hs e que apresentem um preço de fechamentos entre 0.994 e 1.004 em relação a média móvel exponencial de 17 períodos. Após a identificação dos registros com preços afastados da média, é avaliado se o mesmo retornou para próximo da região da média móvel de 17 períodos. Verificando se agora o preço de fechamento encontra-se em uma faixa entre de 0.997 e 1.003, respeitando o limite máximo de horário até as 10:55hs.

Para ser considerado o retorno até as médias, é necessário a identificação inicial de um registro com preço distante da média móvel de 17 períodos. Nos casos onde não for identificado o retorno até as médias pelos critérios de horário e variação de preços, os registros classificados como iniciados com distorções serão desprezados.

4. Análise e Exploração dos Dados

Nessa seção você deve mostrar como foi realizada a análise e exploração dos seus. Mostre as hipóteses levantadas durante essa etapa e os padrões e *insights* identificados.

5. Criação de Modelos de Machine Learning

Conforme o documento de instruções para o TCC, essa etapa não é obrigatória, mas é fortemente recomendada. Caso você crie modelos de *Machine Learn*ing em seu projeto, nessa seção você irá descrever as ferramentas utilizadas. Se você utilizou ferramentas visuais como Knime e Rapid Miner, coloque aqui um print do seu modelo. Caso você tenha escrito scripts em Python, por exemplo, coloque aqui o seu script. Explique as *features* utilizadas, justifique a escolha por determinado modelo, os parâmetros utilizados, etc.

6. Apresentação dos Resultados

Nessa seção você deve apresentar os resultados obtidos. Apresente gráficos, dahsboards, conte a sua história de forma bastante criativa. Aqui você pode utilizar os modelos de Canvas propostos por Dourard (clique [aqui](https://www.louisdorard.com/machine-learning-canvas)) ou por Vasandani (clique [aqui](https://towardsdatascience.com/a-data-science-workflow-canvas-to-kickstart-your-projects-db62556be4d0)).



7. Links

Aqui você deve disponibilizar os links para o vídeo com sua apresentação de 5 minutos e para o repositório contendo os dados utilizados no projeto, scripts criados, etc.

REFERÊNCIAS

Um projeto de Ciência de Dados não requer revisão bibliográfica. Portanto, a inclusão das referências não é obrigatória. No entanto, caso você deseje incluir referências relacionadas às tecnologias ou às metodologias usadas em seu trabalho, relacione-as de acordo com o modelo a seguir.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.

SOBRENOME DO AUTOR, Nome do autor. **Título do livro ou artigo.** Cidade: Editora, ano.