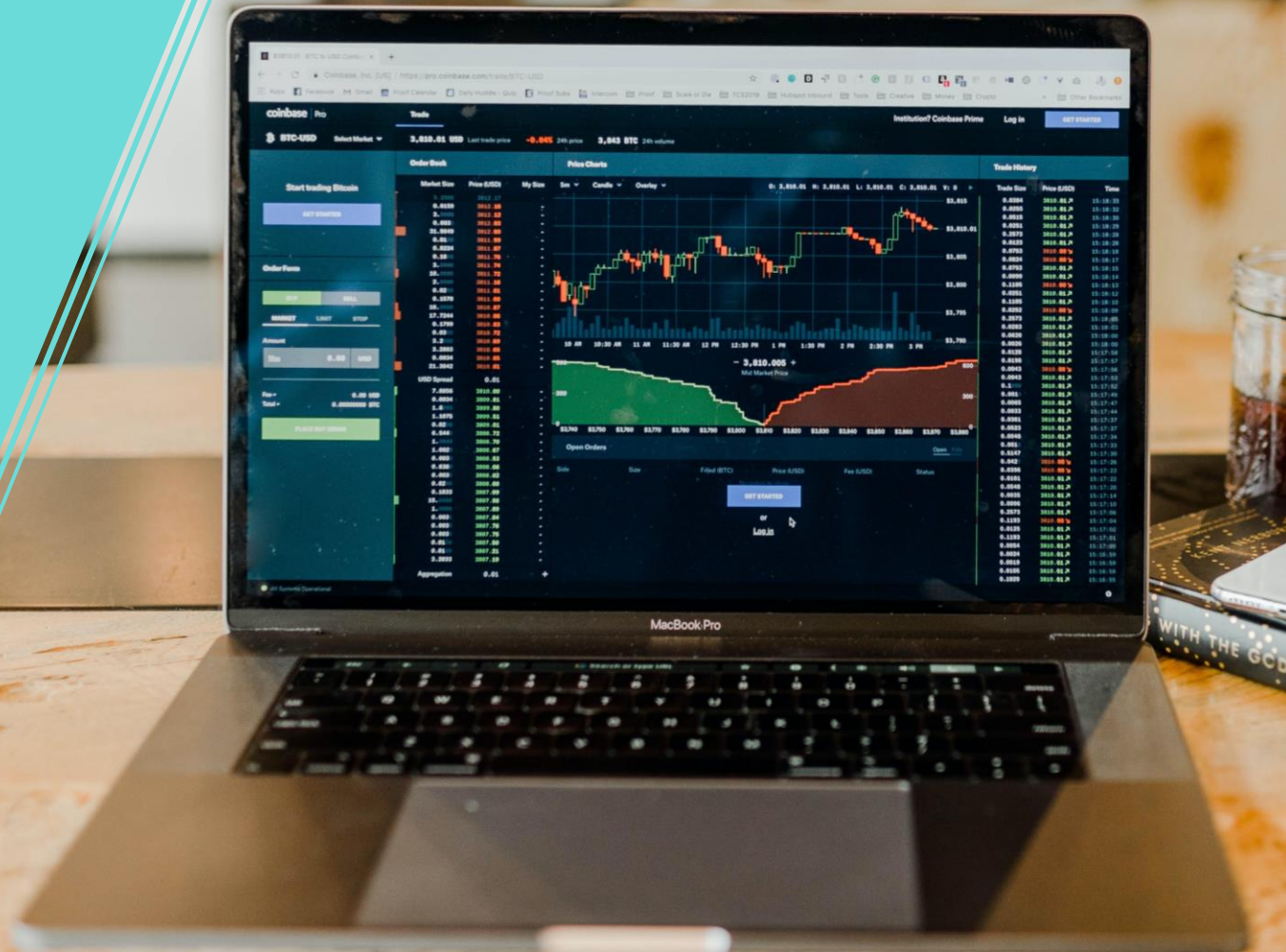


Séries Temporais aplicadas ao Mercado Financeiro

Modelo preditivo de preços das ações da B3

29/05/2020



Pós-Graduação em Análise de Big Data



Aluno:
Caíque Tadeu Filipini

Coordenadores:
Profª Drª Alessandra de Ávila Montini
Prof. Dr. Adolpho Walter Pimazoni Canton



Agenda

- 1. Objetivo
- 2. Contextualização
- 3. Os Dados



1. Objetivo

Desenvolver um modelo que faça a **predição dos preços das ações das empresas listadas na B3** e, com isso, recomendar uma carteira de ativos que supere os resultados da Selic do Ibovespa.

Para isso, serão utilizadas as bases de dados históricos (**séries temporais**) dos preços das ações das empresas. Também é possível que sejam utilizados dados históricos de alguns indicadores macroeconômicos contidos no site do Banco Central do Brasil (IPCA, Selic, dólar, etc.), da performance do Ibovespa, dos dados fundamentalistas das empresas (lucro, endividamento, ROE, etc.) e de distribuições de dividendos e JCP (juros sobre capital próprio).

Para extração dos dados, serão utilizadas técnicas de **Web Scrapping** e requisições por meio de **API** (quando disponibilizada pela fonte).

Para criar o modelo preditivo, serão aplicadas, inicialmente, as **técnicas tradicionais** de séries temporais (ARIMA, SARIMA, etc.), e posteriormente o problema será abordado com **modelos mais robustos e sofisticados** para tratamento de padrões sequenciais, como:

- Discretização da série temporal para aplicação de algoritmos de Machine Learning; e
- Deep Learning (RNN, LSTM, etc.)

2. Contextualização

5

O mercado financeiro, conhecido por sua **volatilidade**, é fortemente influenciado por diversos fatores, vários deles qualitativos e aleatórios, o que dificulta muito a previsão dos preços das ações. Qualquer movimentação política inesperada, por exemplo, poderia causar uma queda brusca generalizada nos ativos da bolsa. Por outro lado, uma notícia de que uma vacina para uma doença crítica teria sido desenvolvida, fortaleceria a confiança dos investidores na economia, implicando diretamente em uma valorização do mercado financeiro. Além do ambiente interno, este setor também é muito **sensível aos indicadores estrangeiros**, principalmente de países como os EUA. Por isso, prever o preço das ações das empresas torna-se uma **tarefa extremamente complexa e desafiadora**.

Contudo, transferindo para os algoritmos apenas a tarefa de encontrar padrões e tendências nos indicadores gerais do mercado financeiro, o ser humano poderia investir seu tempo avaliando outras variáveis que a máquina não consegue capturar. Desta forma, as saídas preditivas do modelo funcionariam não como soberanas e decisivas, mas sim como adicionais às análises do ser humano, **contribuindo e agregando valor na tarefa de compra e venda das ações**.



3. Os Dados

Para o desenvolvimento deste projeto, é necessário um período significativo de dados. Contudo, **datas exageradamente distantes possuem pouca influência no cenário atual**. E a influência que possuem, estão representadas, em grande parte, em dados relativamente mais recentes. Sendo assim, a fim de se obter volume suficiente para treinar, validar e testar os modelos, serão extraídos dados **a partir de 01/01/2010**.

A base de dados principal é composta pelas colunas:

- | | |
|----------|------------------------|
| - Date | - Data da negociação |
| - Open | - Preço de abertura |
| - High | - Maior preço atingido |
| - Low | - Menor preço atingido |
| - Close | - Preço de fechamento |
| - Volume | - Volume negociado |
| - Ticker | - Código da empresa |

Quantidade de ações listadas na B3:	328
Quantidade de dias no período considerado:	2.569
Quantidade máxima* de registros:	842.632

*Como é possível que nem todas as ações listadas na bolsa de valores possuam capital aberto desde o início do período considerado, esta é a quantidade máxima possível de registros que a base geral pode ter, mas não necessariamente a quantidade real.

3. Os Dados

7

Observações:

1. A base descrita anteriormente é a **base principal**. Contudo, ao longo do desenvolvimento do trabalho, é possível que mais bases de dados sejam extraídas a fim de **agregar valor** ao projeto, como por exemplo:

- dados históricos fundamentalistas das empresas
- dados históricos de distribuição de dividendos e JCP
- indicadores macroeconômicos históricos (ex: Selic, IPCA, dólar, etc.)

2. Também é possível, para efeito de simplificação do trabalho e possibilidade de **processamento e armazenamento** em computador tradicional, que se limite a quantidade de ações analisadas de acordo com algum critério.

Séries Temporais aplicadas ao Mercado Financeiro

Modelo preditivo de preços das ações da B3

29/05/2020

