



FIAP





PROJETO INTEGRADO DEEP LEARNING



O DESAFIO:

A **QuantumFinance** está desejando construir um fundo de ações baseado em modelos de Deep Learning. A ideia é desenvolver um modelo que tenha como saída a compra ou venda de uma determinada ação baseado na movimentação do mercado nos últimos 15 dias.

Inicialmente 4 ações foram selecionadas para análise:

- VALE3 – Vale do Rio Doce
- PETR4 – Petrobras
- BBAS3 – Banco do Brasil
- CSNA3 – Companhia Siderúrgica Nacional

O DESAFIO:

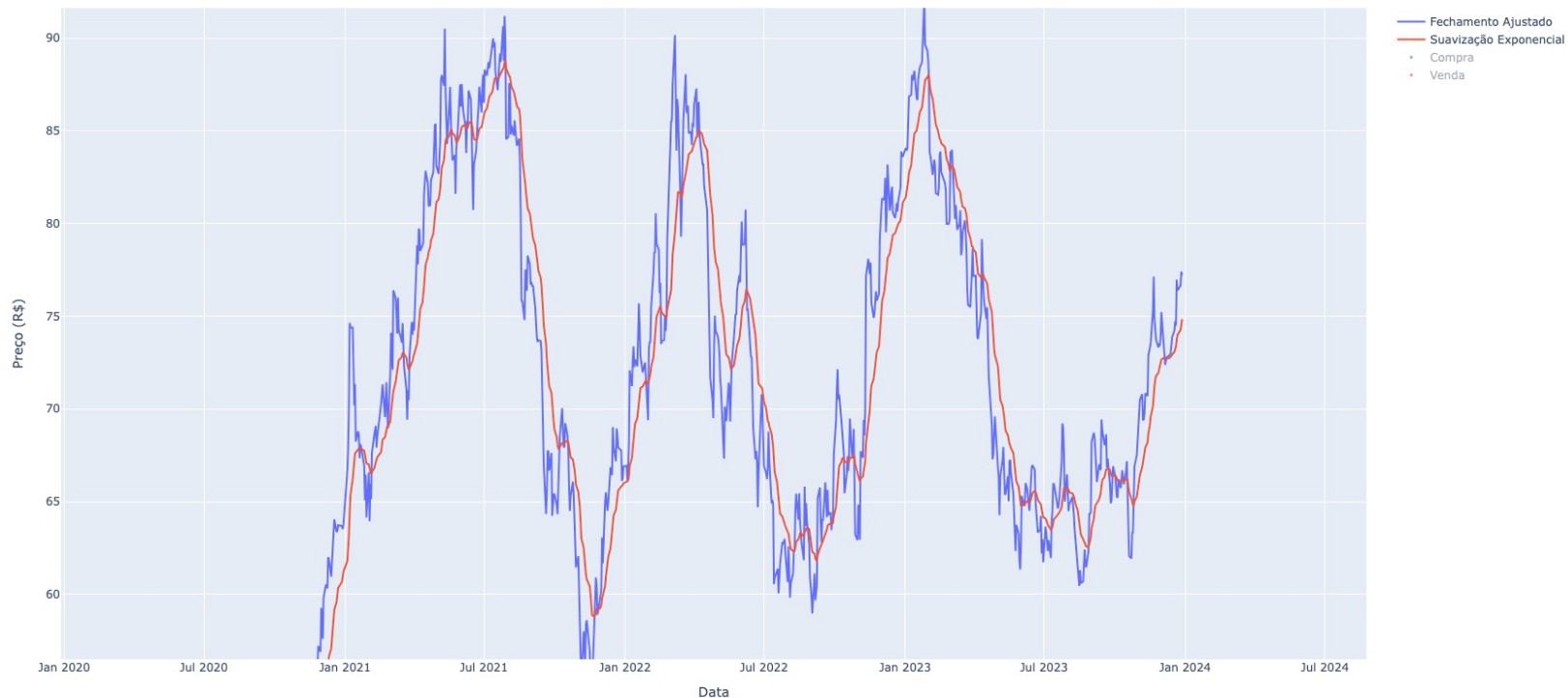
A empresa deseja um modelo de Deep Learning que seja um 'perseguidor de tendencia', ou seja, se o papel sobe o modelo deve comprar, se o papel desce o modelo deve vender.

Para isso a QuantumFinance contratou economistas que fizeram a rotulagem dos dados baseado no valor de fechamento dos ativos. Inicialmente foi feita uma suavização dos preços de fechamento, e por se tratar de um assunto sigiloso, não foi informada a técnica de suavização utilizada. O próximo passo foi realizar a rotulagem dia a dia, para verificar qual devemos comprar e qual devemos vender como segue a imagem a seguir:

O DESAFIO:

Série de preços de fechamento reais e suavizados

Análise Técnica VALE3.SA



O DESAFIO:

Rótulos de compra e venda:


Análise Técnica VALE3.SA





O DESAFIO:

Observe que na imagem anterior que as tendencias de alta estão rotuladas em verde enquanto as tendencias de baixa em vermelho.



Com isso é possível gerar uma predição de compra ou venda para o dia seguinte baseado nos últimos 15 dias de comportamento do mercado.



COMO RESOLVER O PROBLEMA:

Para cada um dos ativos listados será fornecido dois arquivos no formato csv, um para treino e outro para testes. Todos conjuntos de treino começam em janeiro de 2000 e o teste em meados de 2019 até dezembro de 2023.

Também será fornecido um arquivo no formato html mostrando a série de preços, a suavização realizada bem como os rótulos de forma iterativa, como na imagem anterior, apenas para facilitar a interpretação do problema.

Na próxima imagem temos a visualização dos arquivos csv:

COMO RESOLVER O PROBLEMA:

	Date	Close	Smoothed_Close	Label	Past_1_Days_Close	Past_2_Days_Close	Past_3_Days_Close	Past_4_Days_Close	Past_5_Days_Close	Past_6_Days_Close	Past_7_Days_Close
1	2000-01-24	1.062769	1.071645	-1	1.088979	1.079805	1.086357	1.098152	1.069322	1.061459	1.062769
2	2000-01-26	1.087668	1.073247	-1	1.062769	1.088979	1.079805	1.086357	1.098152	1.069322	1.061459
3	2000-01-27	1.064080	1.072330	-1	1.087668	1.062769	1.088979	1.079805	1.086357	1.098152	1.069322
4	2000-01-28	1.036561	1.068753	-1	1.064080	1.087668	1.062769	1.088979	1.079805	1.086357	1.098152
5	2000-01-31	1.035250	1.065403	-1	1.036561	1.064080	1.087668	1.062769	1.088979	1.079805	1.086357
...
4731	2019-03-19	39.001053	38.947137	1	39.852917	40.156639	39.815880	40.119583	39.512161	39.737267	38.555489

Dia da Operação
de Compra/
Venda

Valor do preço
no fim daquele
dia

Valor do preço filtrado no
fim daquele dia (apenas
por curiosidade)

Rótulo (-1 para Venda e +1 para
Compra) para ser operado naquela dia
antes do fechamento (normalmente na
abertura do mercado).

Valor dos preço no dia anterior (D-1) até D-15 que deve
ser usado como feature caso desejado.

COMO RESOLVER O PROBLEMA:

Também será disponibilizado uma pasta contendo um gráfico de barras no formato de imagem png normalizada dos últimos 15 dias de preços para cada dia de treino e teste. Seguem alguns exemplos a seguir:



Imagem 2019-01-18_1.png do conjunto de treino da VALE3, tendencia de alta, recomendação de compra. A última barra (D-1) representa o valor do preço do ativo no dia 2019/01/17.

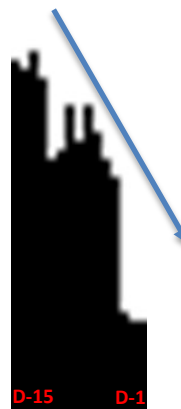


Imagem 2018-11-28_-1.png do conjunto de treino da VALE3, tendencia de baixa, recomendação de venda. A última barra (D-1) representa o valor do preço do ativo no dia 2018/11/27.

COMO RESOLVER O PROBLEMA:

Utilizando os arquivos csv é possível treinar modelos de redes neurais convolucionais 1D (CNN 1D) e redes neurais recorrentes (RNN). Utilizando as imagens é possível treinar redes neurais convolucionais 2D (CNN 2D).

Também é possível extrair mais features dos seguimentos de 15 dias e treinar uma CNN, RNN ou até mesmo uma rede neural mais simplificada.

O QUE É ESPERADO COMO ENTREGA?

A solução **implementada** utilizando modelos de Deep Learning com Tensorflow para as 4 ações e outras soluções **diferenciadas** e **inovadores** (opcional) que possam ser aplicadas ao conjunto de dados proposto.

- Devem ser entregues em formato de Jupyter Notebook;

As quatro ações devem ter modelos treinados, e devem ser apresentados os seguintes indicadores:

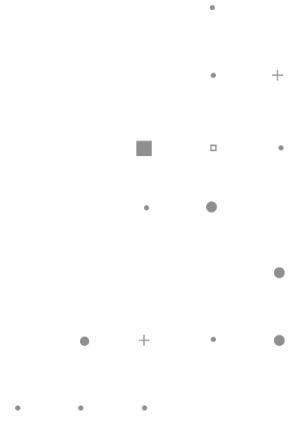
- Acurácia no conjunto de teste de cada modelo;
- Matriz de confusão, precision e recall de cada modelo;
- Desempenho financeiro do modelo via 'backtest' com o objetivo de responder se o modelo gerado daria retorno financeiro (opcional pois é um assunto de finanças, se não fizer não há prejuízo na nota final);



O PROBLEMA:

O que mais seria possível ser feito???


A criatividade é totalmente permitida desde que use modelos de Deep Learning.





CONJUNTO DE DADOS:

Presente no seguinte link:



<https://drive.google.com/drive/folders/1OTU650tmFmeyARV5-IYW7yNacVWPxvOJ?usp=sharing>



OBRIGADO

FIAP

Copyright © 2024 | Professor Felipe Gustavo Silva Teodoro

Todos os direitos reservados. A reprodução ou divulgação total ou parcial deste documento é expressamente proibida sem consentimento formal, por escrito, do professor(a)/autor(a).



FIAP

