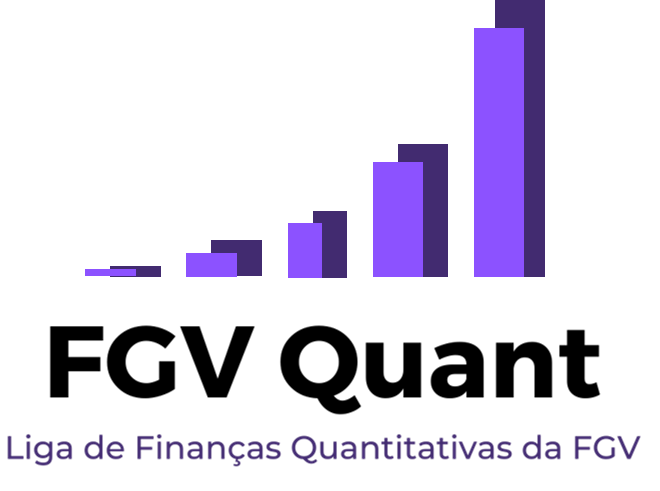
****

**FGV Quant**

**Liga de Finanças Quantitativas da FGV**

Hugo Vinícius Ferreira Azevedo

**MODELOS DE MACHINE E DEEP LEARNING APLICADOS NA PREVISÃO DE AÇÕES BRASILEIRAS**

**SÃO PAULO**

**2021**

**1. DESCRIÇÃO DO PROJETO**

No contexto da Hipótese do Mercado Adaptativo, que afirma que os preços das ações são previsíveis, a análise de previsão utilizando modelos de Machine Learning, tais como Random Forest (RF), Artificial Neural Network (ANN) e Support Vector Machine (SVM) tem sido amplamente utilizada. Entretanto, estes modelos não puderam satisfazer os requisitos de desempenho em cenários parciais com flutuações complexas. Portanto, este trabalho pretende desenvolver uma implementação em ambiente Python para realizar previsão de ações usando ANN, SVM e também um modelo de aprendizado profundo chamado Long-Short-Term-Memory Network (LSTM), bem como seus híbridos acoplados à Transformada Discreta de Wavelet (DWT) como técnica de pré-processamento dos dados. Os resultados obtidos serão comparados ao desempenho de um algoritmo de RF, que é amplamente utilizado como um modelo de base.

Os dados diários das ações ITUB4, ABEV3 e KLBN11 foram coletados de 2016 a 2021 pela yahoo finance e a validação da janela deslizante foi usada para avaliar o desempenho do prognóstico.

Espera-se que os resultados indiquem que o modelo híbrido de aprendizagem profunda teve um melhor desempenho quando comparado com os modelos de aprendizagem de máquinas híbridas e não-híbridas e o uso de indicadores técnicos econômicos como recursos de entrada será discutido.

**2. CRONOGRAMA**

O trabalho será desenvolvido a partir das atividades descritas a seguir:

1. Montar o dataset
2. Análise exploratória dos dados
3. Filtrar e fazer o pré-processamento
4. Testar modelos de aprendizagem de máquina (SVM, RF)
5. Testar modelos de aprendizagem profunda (LSTM, ANN)
6. Testar modelos híbridos (acoplados à DWT)
7. Avaliar modelos e discutir resultados
8. Fazer backtesting (Validation set)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ATIVIDADES | Semanas | | | | | | | | |
| Junho | | | | | Julho | | | |
| 1ª | 2ª | 3ª | 4ª | 5ª | 1ª | 2ª | 3ª | 4ª |
| 1 | X | X |  |  |  |  |  |  |  |
| 2 |  | X |  |  |  |  |  |  |  |
| 3 |  |  | X |  |  |  |  |  |  |
| 4 |  |  |  | X |  |  |  |  |  |
| 5 |  |  |  |  | X |  |  |  |  |
| 6 |  |  |  |  |  | X | X |  |  |
| 7 |  |  |  |  |  |  |  | X |  |
| 8 |  |  |  |  |  |  |  |  | X |

**3. ÍNDICES TÉCNICOS**

Alguns indicadores técnicos serão calculados para complementarem as features dos modelos de machine e deep learning. São eles:

* 1. **Índice de força relativa (Relative Strength Index – RSI)**
  2. **Oscilador estocástico (Stochastic Oscillator - %K)**

Onde,

= Preço atual de fechamento;

= Menor preço nos últimos 14 dias;

= Maior preço nos últimos 14 dias.

* 1. **Oscilador de Larry Williams (%R)**
  2. **Taxa de mudança dos preços (Price Rate of Change – PROC)**

Onde,

= Taxa de mudança do preço no tempo ;

= Preço de fechamento no tempo .

* 1. **Indicador Saldo de Volume (On-Balance Volume - OBV)**

Onde,

= Indicador saldo de volume no tempo ;

= Volume negociado no tempo ;

= Preço de fechamento no tempo .

* 1. **Convergência e divergência da média móvel (Moving Average Convergence Divergence – MACD)**

Onde,

= Série temporal dos dados de fechamento;

= Média exponencial de dias.