

Name: Click or tap here to enter text.

Tema: Artificial Intelligence Application

Enunciado:

1. Contexto do Caso

O fundo de investimentos “Agile Capital” quer melhorar sua estratégia de alocação dinâmica, prevendo preços diários de Bitcoin (BTC-USD) e de uma ação blue-chip (ex.: Apple – AAPL), com horizonte de 14 dias. O objetivo é testar sinais de trade baseados em forecast e avaliar performance relativa entre classes de ativos.

2. Objetivos de Aprendizagem

- Coletar dados financeiros de múltiplas fontes (Kaggle e API do Yahoo Finance).
- Construir pipeline de séries temporais com janelas e variáveis exógenas (volume, indicadores técnicos).
- Treinar RNNs (SimpleRNN, LSTM, GRU) e comparar com modelos tradicionais (MS Excel).
- Avaliar métricas de forecast: RMSE, MAE, MAPE e análise de retornos (Sharpe ratio simulado).
- Debater limites de RNNs em dados financeiros voláteis e falar sobre risco de overfitting.

3. Descrição dos Datasets

- Bitcoin Historical Data (Kaggle): preços diários de BTC-USD desde 2014, com Open/High/Low/Close/Volume.
- Stocks Historical Data (Kaggle ou Yahoo Finance): preço diário de AAPL (ou outra blue-chip) no mesmo período.
- Indicadores Técnicos (gerados): médias móveis (SMA, EMA), RSI, MACD.
- Formato: CSV único por ativo; total de ~3.000 registros (2014–2024) para cada um.

4. Ferramentas e Tecnologias

- Ambiente: Google Colab / Kaggle Notebooks.
- Bibliotecas: pandas, numpy, ta (technical analysis)
- scikit-learn (preprocess)
- TensorFlow/Keras ou PyTorch (RNNs)
- matplotlib/plotly (visualização)

- Deploy opcional: Streamlit para dashboard interativo.

5. Etapas Detalhadas

1. Coleta & Ingestão
Baixar CSVs do Kaggle + usar yfinance para garantir sincronização de datas.
2. EDA & Feature Eng. - Plotar séries de preço e volume.
3. Calcular SMA, EMA, RSI, MACD e adicionar como features exógenas.
4. Pré-processamento - Normalizar com MinMax ou StandardScaler (fit no train).
5. Tratar datas faltantes e feriados (forward fill).
6. RNNs: SimpleRNN, LSTM e GRU com 1–2 camadas, dropout e estado retornável (stateful).
7. Treino & Validação
8. TimeSeriesSplit ou validação em rolling windows.
9. EarlyStopping e redução de LR.
10. Avaliação & Comparação
11. RMSE, MAE, MAPE para cada ativo e modelo.
12. Simular retorno hipotético: buy-and-hold vs. forecast-based signals (long-only).
13. Análise de Risco
14. Calcular Sharpe ratio simulado.
15. Deploy & Dashboard
16. (Opcional) Criar app em Streamlit mostrando forecast para BTC e AAPL, indicadores e performance.

6. Questões para Debate

- Volatilidade vs. dependência temporal: RNNs conseguem capturar picos e crashes em criptomoedas?
- Exógenas x endógenas: até que ponto indicadores técnicos ajudam no forecast?
- Modelos tradicionais vs. deep learning: onde modelos clássicos ainda ganham?
- Gerenciamento de risco: usar forecast para trade real é viável? Que frações alocar?
- Sobretreinamento: como detectar e evitar overfitting em dados financeiros ruidosos?

7. Entregáveis Esperados

- Notebook (.ipynb) bem estruturado: coleta, EDA, modelagem e avaliação.
- Relatório (8–10 páginas): metodologia, comparativo de modelos e simulação de estratégia. Insights principais e recomendações de uso em carteira.
- Dashboard Streamlit (opcional): forecast interativo para BTC e AAPL.

8. Cronograma recomendado

- Coleta de dados, EDA e indicadores - Gráficos de séries e indicadores
- Dataset pronto para treino
- Treino de modelos e tuning - Comparativo de performance
- Simulação de estratégia e deploy - Relatório final + dashboard

Especificações quanto ao conteúdo:

- No seu referencial teórico, traga os autores renomados na área e publicações recentes – até 5 anos.
- Sob essas perspectivas de análise, você deve construir um texto que procure responder às indagações do professor.
- Lembre-se, um artigo científico não se trata de um texto opiniático, sem embasamento teórico ou referências bibliográficas; ou seja, obrigatoriamente para você ser bem avaliado, deverá construir um texto com base no rigor da pesquisa científica e portanto, atendendo a todos os critérios metodológicos a seguir, descritos.

Forma metodológica: ESTUDO DE CASO

1. Escrita científica na qual o texto esteja embasado em autores, pesquisadores e organizações relevantes da área publicações recentes (até 5 anos). Os textos devem conter, o mínimo de 5 autores citados; e pode-se aceitar, um ou dois autores mais antigos que ultrapassem 5 anos.
2. Todos os textos devem trazer no final, obrigatoriamente, as referências bibliográficas completas que forem citadas dentro do texto; e devem conter também, as referências gerais, de inspiração do estudante.
3. Os textos devem apresentar título, objetivos e conclusões ou considerações finais. Caso se trate de um artigo de análise ou de conclusão de um experimento, deve ser destacada a metodologia utilizada.
4. Os textos não podem apresentar similaridade externa e interna, ou seja, não podem ser copiados entre os próprios estudantes; e nem podem ser cópia de terceiros, o que inclui materiais gerados pela inteligência artificial; sob o risco de zerar a nota.
5. O texto deve obedecer às regras de concordância verbal e nominal, além de correção ortográfica.
6. O texto deve apresentar um mínimo de 10 mil caracteres e um máximo de 15 mil caracteres (com espaços). Obs.: Cuidar para não trazer páginas “a mais” de referências bibliográficas, do que de texto propriamente construído, para suprir o quantitativo de caracteres, pois isso será penalizado.
7. Deve ser conciso, objetivo, fluido e principalmente, autoral (embora, e obrigatoriamente embasado nos autores estudados)