

# Comparação de Desempenho entre Redes Neurais LSTM e GRU na Previsão de Preço de Ações Utilizando PyTorch no Google Colab

Eduardo Junqueira - Engenharia de Software Puc Minas

## I. RESUMO

Este estudo realiza uma comparação entre as arquiteturas de redes neurais recorrentes LSTM (Long Short-Term Memory) e GRU (Gated Recurrent Unit), aplicadas à tarefa de previsão dos preços de fechamento diário de ações de grandes empresas de tecnologia: Apple, Microsoft, Alphabet, Meta e Nvidia. Os dados foram coletados automaticamente por meio da biblioteca yfinance, e os modelos foram implementados em PyTorch e treinados no ambiente Google Colab. A GRU apresentou desempenho ligeiramente superior em termos de erro quadrático médio ( $MSE = 0.0273$ ), superando a LSTM ( $MSE = 0.0314$ ) com menor tempo de treinamento. O estudo conclui que a GRU pode ser mais eficiente para aplicações de previsão de séries temporais financeiras em curto prazo.

## II. INTRODUÇÃO

A previsão do comportamento de ativos financeiros é uma tarefa desafiadora, mas de grande valor prático, especialmente em contextos de suporte à decisão no mercado financeiro. Modelos de redes neurais recorrentes (RNNs) são comumente utilizados em tarefas que envolvem séries temporais. As variantes LSTM(1997) e GRU(2014) são projetadas para capturar padrões temporais de longo prazo, sendo particularmente adequadas para dados financeiros. Este trabalho visa comparar essas duas arquiteturas sob condições controladas e com dados reais do mercado, com o objetivo de avaliar o desempenho preditivo e a eficiência computacional de cada uma.

A previsão de séries temporais é uma área crítica para diversos setores, como energia, saúde, clima e, especialmente, finanças. No contexto dos mercados financeiros, prever o comportamento futuro de ativos com base em dados históricos pode proporcionar vantagens competitivas, auxiliar na tomada de decisão e na gestão de riscos.

Nos últimos anos, redes neurais recorrentes (RNNs) ganharam destaque por sua capacidade de lidar com dependências temporais nos dados. No entanto, as RNNs tradicionais sofrem com o problema do desaparecimento do gradiente, o que compromete sua eficiência em tarefas que exigem memória de longo prazo. Para contornar esse problema, surgiram variantes mais sofisticadas, como a LSTM e a GRU.

A LSTM introduz estruturas de memória e portas de controle que permitem à rede aprender quais informações devem ser lembradas ou esquecidas. Já a GRU, uma versão mais

simples e computacionalmente eficiente, utiliza menos portas, mantendo a capacidade de capturar dependências temporais.

Este estudo tem como foco comparar, de forma empírica, o desempenho dessas duas arquiteturas em um problema real de previsão financeira, utilizando dados de mercado atualizados, extraídos automaticamente da internet, e explorando ferramentas modernas de aprendizado profundo.

## III. MÉTODO

Os dados foram coletados automaticamente com o uso da biblioteca yfinance, da seguinte forma. Essa rotina extraiu os preços de fechamento diário das cinco ações para o período de 1 ano, salvando os dados em um arquivo CSV que posteriormente foi carregado no Google Colab.

Preparação dos Dados: Normalização dos dados com Min-MaxScaler. Transformação em janelas de 30 dias para prever o próximo valor (estrutura seq2one). Conversão dos dados para tensores e separação em conjuntos de treinamento 80 por cento e teste 20 por cento.

Com os preços de fechamento diários das ações da Apple, Microsoft, Meta, Alphabet e Nvidia no período de 1 ano. A coluna Close foi selecionada por representar o valor de fechamento de cada ação em um dado dia, sendo a mais utilizada em análises técnicas. Após a coleta, os dados foram normalizados utilizando MinMaxScaler, de forma a restringir todos os valores ao intervalo 0 a 1, o que facilita o aprendizado das redes neurais. Em seguida, foram organizados em janelas deslizantes de 30 dias: a cada sequência de 30 dias foi associada a previsão do próximo dia.

Os dados foram convertidos em tensores do PyTorch e divididos em dois conjuntos: treinamento 20 dos dados, teste 80 restantes, utilizados exclusivamente para avaliar o modelo. Os modelos utilizados foram implementados e treinados dois modelos distintos:

1. LSTM (Long Short-Term Memory) A arquitetura LSTM utilizada possui uma camada oculta e uma camada linear na saída. Foi projetada para reter informações por períodos mais longos por meio de mecanismos de portas, como as portas de entrada, esquecimento e saída.

2. GRU (Gated Recurrent Unit) A GRU, por sua vez, é uma versão mais leve da LSTM. Utiliza menos portas, sendo assim mais rápida de treinar e menos exigente em termos computacionais. O modelo adotado também possui uma camada oculta e uma camada linear para projeção da saída.

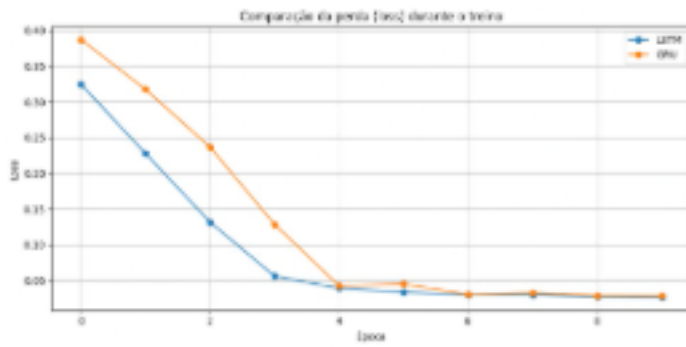


Fig. 1. Figura 1 - Resultados

O treinamento foi conduzido com os seguintes parâmetros: Função de perda: Erro Quadrático Médio Épocas: 10 Batch size: 32

Durante o treinamento, a perda média por época foi registrada. Após a conclusão, os modelos foram avaliados com base no erro de teste (MSE), para verificar a capacidade de generalização.

#### IV. RESULTADOS

Ambas as arquiteturas foram eficazes na tarefa de previsão de séries temporais, apresentando erros baixos com apenas 10 épocas de treinamento. No entanto, a **GRU se destacou** com um menor erro de teste e menor tempo de convergência, o que a torna uma alternativa mais eficiente para aplicações de curto prazo em finanças. Embora a LSTM continue sendo uma arquitetura robusta e confiável, a GRU provou ser uma escolha sólida e leve para esse cenário específico. Resultados

LSTM – Resultados de Treinamento e Teste Epoch 10/10 - Loss: 0.028569 Loss no teste: 0.031427 GRU – Resultados de Treinamento e Teste Epoch 10/10 - Loss: 0.026366 Loss no teste: 0.027301

Ambas as arquiteturas demonstraram ser capazes de capturar padrões nos dados e realizar previsões razoavelmente precisas. No entanto, a GRU apresentou menor erro no conjunto de teste, o que indica melhor generalização.

Além disso, a GRU foi treinada de forma mais rápida, reforçando sua vantagem quando se consideram recursos computacionais limitados ou aplicações com restrições de tempo. Os resultados deste estudo indicam que, para a tarefa de previsão de preços de fechamento de ações com séries temporais relativamente curtas (1 ano de dados), tanto LSTM quanto GRU são opções viáveis e eficientes. Entretanto, a GRU se mostrou levemente superior, tanto em termos de precisão quanto em tempo de treinamento. Sua estrutura simplificada e menor demanda computacional tornam-na ideal para aplicações em tempo real ou que exigem respostas rápidas, com operações financeiras automatizadas. Vale destacar que a LSTM ainda pode ser vantajosa em séries temporais mais longas ou com padrões sazonais mais complexos, onde a retenção de informações de longo prazo é mais crítica.

#### V. CONCLUSÕES

Com base nos resultados apresentados, é possível tirar algumas conclusões importantes sobre o desempenho comparativo entre as redes neurais LSTM e GRU na tarefa de previsão de preços de ações. Tanto a LSTM quanto a GRU conseguiram aprender padrões dos dados históricos de preços de ações e generalizar para novos dados (teste), com valores relativamente baixos de erro médio quadrático (MSE). Isso mostra que RNNs avançadas são apropriadas para séries temporais financeiras.

A GRU obteve menor erro tanto no treinamento quanto no teste, indicando que: Aprendeu melhor os padrões dos dados. Generalizou melhor para dados que não viu durante o treinamento. GRU pode ser preferível em muitos casos. Isso a torna uma boa escolha para aplicações com recursos limitados como notebooks no Google Colab ou quando se busca agilidade no desenvolvimento. Apesar de a GRU ter tido melhor desempenho, a diferença não é extremamente significativa. Em problemas mais complexos ou com mais dados, a LSTM pode eventualmente ter vantagem pela sua maior capacidade de modelar dependências temporais mais longas.

Neste experimento, a GRU superou ligeiramente a LSTM em termos de erro de previsão e eficiência. Portanto, para tarefas similares de previsão de séries temporais financeiras com dados moderados, a GRU pode ser uma escolha mais prática e eficaz.

#### REFERENCES

- [1] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
- [2] Cho, Kyunghyun, et al. "Learning phrase representations using RNN encoder-decoder for statistical machine translation." arXiv preprint arXiv:1406.1078 (2014).