

Trabalho em Redes neurais profundas recorrentes

1st João Pedro Aguiar Formiga Matos *EMC-UFG*
Universidade Federal de Goiás
Goiânia, Brasil
Joao.formiga@discente.ufg.br

Abstract—A análise e predição de séries temporais representam um desafio significativo, que tem sido progressivamente abordado com o advento das redes neurais, apesar de sua complexidade inerente. Este estudo propõe a implementação e avaliação de três modelos de predição de séries temporais - MLP, LSTM e GRU - cada um com suas características distintas. O objetivo principal é treinar estas redes para prever séries temporais, buscando minimizar o erro em relação à série real, preservando suas características intrínsecas. Este artigo é o resultado de um estudo conduzido como parte da disciplina de Redes Neurais Profundas (2023.1) na Universidade Federal de Goiás, Faculdade das Engenharias, com o foco principal na avaliação do desempenho dessas três redes neurais na predição de uma série de Mackey-Glas. O restante do artigo detalha a metodologia adotada, a implementação do modelo, os experimentos realizados, os resultados obtidos e as considerações finais.

I. METODOLOGIA

A. Introdução

Este trabalho tem como objetivo o desenvolvimento e a avaliação de três modelos de redes neurais recorrentes, com foco na predição de uma série temporal complexa, especificamente a curva de Mackey-Glas. A escolha desta série temporal se deve à sua natureza não-linear e caótica, que apresenta um desafio significativo para os métodos de predição. O principal objetivo é minimizar o erro de predição, utilizando três métricas distintas para avaliar a precisão dos modelos: o Erro Quadrático Médio (EQM), o Teste de Kolmogorov-Smirnov e os Erros Quadráticos Médios Normalizados (EQMN1 e EQMN2). Estas métricas são calculadas com base nas seguintes fórmulas:

$$EQM = E[(\hat{x} - x)^2] \quad (1)$$

$$EQMN1 = \frac{E[(\hat{x} - x)^2]}{E[(\mu - x)^2]} \quad (2)$$

$$EQMN2 = \frac{E[(\hat{x} - x)^2]}{E[(\hat{x}_{pa} - x)^2]} \quad (3)$$

A escolha destas métricas específicas se deve à sua capacidade de fornecer uma avaliação abrangente do desempenho dos modelos, considerando tanto a precisão da predição quanto a distribuição dos erros. Através deste estudo, buscou-se não apenas desenvolver modelos de redes neurais eficazes para a predição de séries temporais, mas também contribuir para a

compreensão dos desafios e das melhores práticas nesta área de pesquisa em constante evolução.

Este trabalho começou com a definição dos modelos e a preparação dos dados, que foram importados para o arquivo principal. Em seguida, as funções necessárias foram definidas e o código foi executado. Durante o desenvolvimento do projeto, observou-se que os algoritmos eram executados de maneira eficiente na CPU do computador, tornando desnecessário o uso de uma GPU ou a migração para uma plataforma de computação em nuvem, como o Google Colab-oratory.

As arquiteturas dos três modelos de redes neurais recorrentes propostos são caracterizadas por sua simplicidade, consistindo de apenas duas camadas. Esta escolha de design foi motivada pela necessidade de equilibrar a complexidade do modelo e a capacidade computacional disponível, bem como pela intenção de facilitar a interpretação dos resultados.

Cada uma das camadas desempenha um papel específico na transformação dos dados de entrada em previsões de saída. A primeira camada é responsável por extrair características relevantes dos dados de entrada, enquanto a segunda camada utiliza essas características para gerar a previsão.

Ao longo da execução deste trabalho, enfrentou-se desafios relacionados ao dimensionamento de tensores e ao tratamento de dados. Estes problemas foram resolvidos apenas na fase final do projeto, o que impactou o tempo disponível para aprimoramento dos modelos e experimentação com diferentes arquiteturas.

Um ponto crítico foi a falta de pré-processamento dos dados provenientes da série de Mackey-Glas. Esta omissão resultou em um esforço considerável para modificar e adaptar os dados para uso nos modelos, consumindo tempo que poderia ter sido dedicado à otimização dos modelos e à experimentação com diferentes arquiteturas de redes neurais.

Essa experiência ressalta a importância do tratamento adequado dos dados em projetos de aprendizado de máquina. A preparação dos dados é uma etapa crucial que pode influenciar significativamente a eficiência do desenvolvimento do modelo e a qualidade dos resultados obtidos. No futuro, planejo dedicar mais atenção a esta etapa, para maximizar o tempo disponível para a experimentação e otimização dos modelos.

Na próxima seção abaixo serão mostrados os dados obtidos pelas redes neurais e as conclusões obtidas de cada uma.

B. MLP

Dados do modelo MLP:

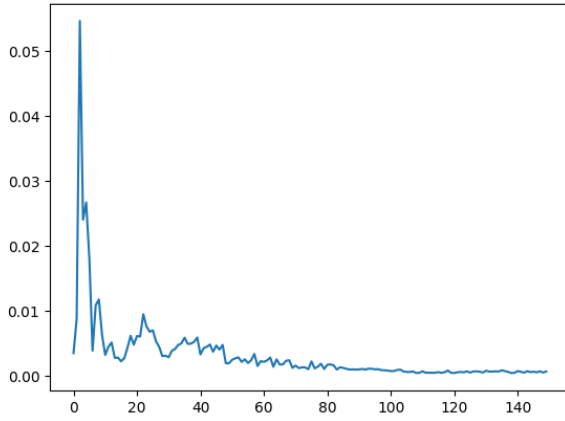


Fig. 1. loss mlp

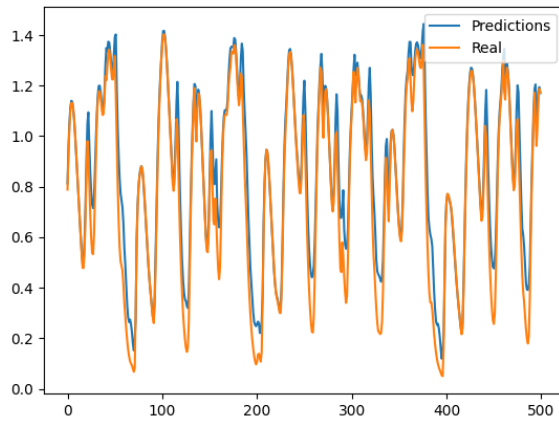


Fig. 2. resultados preditos MLP

TABLE I
RESULTADOS DA MLP PARA OS DADOS DE TREINO.

EQM	EQMN1	EQMN2	pvalue
0.00350	0.02640	0.40466	0.45992

C. LSTM

Dados da LSTM:

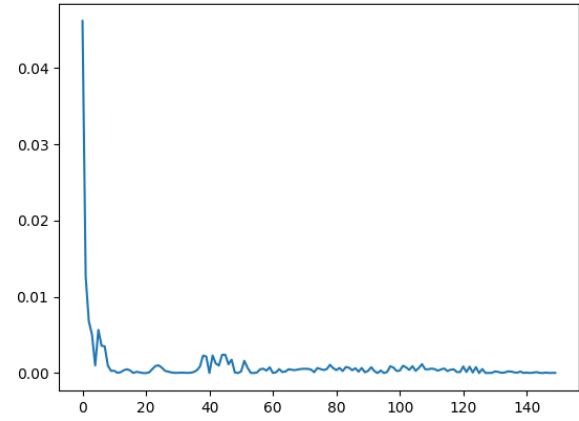


Fig. 3. loss LSTM

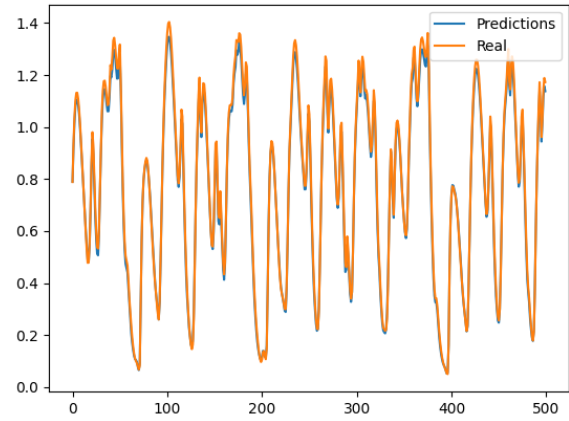


Fig. 4. resultados preditos LSTM

TABLE II
RESULTADOS DA LSTM PARA OS DADOS DE TREINO.

EQM	EQMN1	EQMN2	pvalue
0.000631339	0.004761029	0.072977157	0.56002

II. GRU

Dados da GRU:

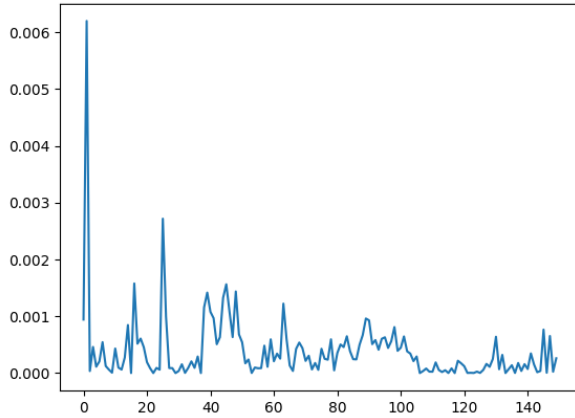


Fig. 5. loss GRU

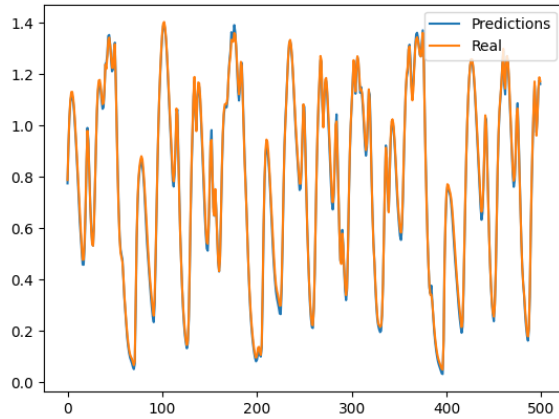


Fig. 6. resultados preditos GRU

TABLE III
RESULTADOS DA GRU PARA OS DADOS DE TREINO.

EQM	EQMN1	EQMN2	pvalue
0.000309964	0.002337488	0.035829081	0.99595

A. Conclusões

Com base nos resultados obtidos, podemos concluir que, considerando o número de épocas e camadas ocultas utilizadas, o modelo GRU apresentou a melhor performance, conforme avaliado pelo teste de Kolmogorov. No entanto, é importante destacar que os modelos LSTM e MLP também demonstraram um desempenho notável, mesmo com um número limitado de pontos. Isso sugere que, apesar de serem considerados modelos mais "simples", eles possuem uma capacidade significativa de aprendizado e predição.

Acredita-se que o desempenho desses modelos possa ser ainda mais aprimorado com o aumento do número de camadas e épocas de treinamento. Essa hipótese é sustentada pela natureza das redes neurais, que tendem a melhorar sua capacidade de generalização à medida que são expostas a mais dados e têm a oportunidade de ajustar seus pesos através de mais iterações de treinamento. Portanto, os resultados promissores obtidos neste estudo servem como um incentivo para futuras investigações, que podem explorar a otimização desses modelos através do ajuste de seus hiperparâmetros.