

Exercício 4

Matheus Arataque Uema - 10276949

Herlisson Maciel Bezerra - 10224760

Dhyogo Nunes Costa - 13096109

SCC5809 - Redes Neurais e Aprendizado Profundo

Resumo

Este relatório apresenta a aplicação de quatro modelos de Redes Neurais Recorrentes (RNNs) para a previsão da variação mensal do Índice de Preços ao Consumidor Amplo (IPCA), utilizando a série temporal de 2023. Os modelos analisados incluem: RNN com função de ativação *TANH*, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU) e Temporal Convolutional Network (TCN). Todos os modelos foram treinados utilizando a técnica de validação cruzada específica para séries temporais, *TimeSeriesSplit*, e avaliados com base na métrica de erro quadrático médio da raiz (RMSE). Os resultados indicam variações significativas de desempenho entre os diferentes modelos.

1 Introdução

A previsão de séries temporais é um problema fundamental em diversos campos, como finanças, economia e climatologia. Neste trabalho, aplicamos quatro arquiteturas populares de Redes Neurais Recorrentes (RNNs) para prever a variação mensal do IPCA, um indicador importante da inflação no Brasil.

As redes RNNs são adequadas para dados sequenciais devido à sua capacidade de manter informações ao longo do tempo. As variantes LSTM e GRU são especialmente projetadas para lidar com dependências de longo prazo, enquanto o TCN é uma abordagem convolucional que tem sido usada com sucesso em problemas de séries temporais.

2 Metodologia

2.1 Dados

Os dados utilizados correspondem à variação mensal do IPCA em 2023, disponíveis no arquivo `inflacao.csv`. O conjunto de dados foi normalizado utilizando o escalador `MinMaxScaler`, de forma a garantir que os valores permanecessem no intervalo $[0, 1]$. A implementação do código pode ser visualizada no

repositório criado no *GitHub*¹, onde também contém instruções para a execução do código.

2.2 Modelos

Os quatro modelos usados neste experimento foram:

- **RNN**: Uma rede recorrente simples, com função de ativação *tanh* e uma única camada recorrente de 50 neurônios.
- **LSTM**: Uma rede do tipo Long Short-Term Memory, com uma camada de 50 neurônios.
- **GRU**: Uma rede Gated Recurrent Unit, também com 50 neurônios.
- **TCN**: Uma rede convolucional temporal, com duas camadas convolucionais e um mecanismo de conexão residual.

Todos os modelos foram configurados com 50 neurônios e treinados pelo mesmo número de épocas (até 100) para garantir uma comparação justa.

2.3 Treinamento e Validação

Utilizamos a técnica de validação cruzada específica para séries temporais, *TimeSeriesSplit*, disponível no `sklearn.model.selection`, com 5 divisões (*splits*). Em cada divisão, o conjunto de dados foi dividido em treino e teste de acordo com o tempo, respeitando a natureza sequencial dos dados. A métrica de desempenho utilizada para avaliar os modelos foi o *Root Mean Squared Error* (RMSE).

O treinamento de cada modelo foi realizado com **EarlyStopping** para evitar o sobreajuste, monitorando a perda (*loss*) no conjunto de treino.

3 Resultados

A Tabela 1 apresenta os valores de RMSE obtidos para cada modelo após a validação cruzada.

O modelo TANH RNN apresentou o melhor desempenho em termos de RMSE, seguido pelo GRU e LSTM. O modelo TCN, embora popular em alguns domínios de séries temporais, apresentou o pior desempenho nesta tarefa específica.

¹Repositório do Código Disponível em: <https://github.com/MatheusUema/neural-networks-study>. Acesso em: 27 de ago. de 2024.

Tabela 1: Resultados de RMSE para os Modelos Avaliados

Modelo	RMSE Médio
TANH RNN	0.0369
LSTM	0.0556
GRU	0.0493
TCN	0.0838

4 Discussão

Os resultados indicam que, para o conjunto de dados utilizado, a arquitetura simples de RNN com função de ativação *tanh* foi a mais eficaz, enquanto o TCN apresentou o desempenho mais fraco. Embora o LSTM e o GRU sejam frequentemente citados como superiores em tarefas de séries temporais devido à sua capacidade de capturar dependências de longo prazo, o desempenho mais simples do RNN pode indicar que a série de preços ao consumidor utilizada não requer modelagem de dependências tão complexas.

5 Conclusão

Neste estudo, comparamos quatro modelos de redes neurais aplicados à previsão do IPCA. A arquitetura TANH RNN obteve o menor RMSE, sugerindo que, para este conjunto de dados, um modelo mais simples pode ser preferível. No entanto, mais experimentos com diferentes hiperparâmetros e arquiteturas podem ser necessários para validar esses achados.