**Comparando Modelos LSTM e GRU para Prever a Condição de uma Prensa de Papel Pulp**

**Autores:**

* Balduíno César Mateus¹²\*
* Mateus Mendes³⁴
* José Torres Farinha⁴⁵
* Rui Assis¹
* António Marques Cardoso²

**Resumo:**

A precisão de um sistema preditivo é crucial para a manutenção preditiva e para apoiar decisões corretas nos momentos certos. Modelos estatísticos, como ARIMA e SARIMA, são incapazes de descrever a natureza estocástica dos dados. Redes neurais, como **memória de longo curto prazo** (LSTM) e **unidade recorrente com porta** (GRU), são bons preditores para dados univariados e multivariados. O presente artigo descreve um estudo de caso onde são comparados os desempenhos das memórias de longo curto prazo e das unidades recorrentes com porta, com base em diferentes hiperparâmetros. Em geral, as unidades recorrentes com porta exibem melhor desempenho, baseado em um estudo de caso sobre prensas de papel pulp. O resultado final demonstra que, para maximizar a disponibilidade do equipamento, as unidades recorrentes com porta, como demonstrado no artigo, são as melhores opções.

**Palavras-chave:**

* LSTM
* Rede neural recorrente
* GRU
* Prensa de papel
* Manutenção preditiva

**Conteúdo do Artigo:**

**1. Introdução:**

A manutenção preditiva é essencial para garantir a disponibilidade e eficiência de equipamentos industriais, como prensas de papel pulp. Modelos estatísticos tradicionais, como ARIMA e SARIMA, têm limitações em lidar com a natureza estocástica dos dados. Redes neurais recorrentes (RNNs), especificamente LSTM e GRU, são alternativas poderosas devido à sua capacidade de capturar dependências de longo prazo em séries temporais.

**2. Modelos Utilizados:**

* **LSTM (Long Short-Term Memory):**
  + Capaz de lembrar informações por longos períodos e é adequado para dados sequenciais.
  + Usa "gates" (portas) para regular o fluxo de informações.
* **GRU (Gated Recurrent Unit):**
  + Variante mais simples e eficiente do LSTM.
  + **Possui menos parâmetros e pode ser treinada mais rapidamente.**

**3. Metodologia:**

Foram realizados experimentos comparativos entre LSTM e GRU com diferentes configurações de hiperparâmetros, como número de camadas e neurônios por camada. Os dados utilizados no estudo provêm de prensas de papel pulp, focando em prever falhas e a necessidade de manutenção.

**4. Resultados:**

Os resultados indicam que, em geral, os modelos GRU superam os modelos LSTM em termos de precisão e eficiência. A simplicidade e a menor quantidade de parâmetros do GRU contribuem para um treinamento mais rápido e um desempenho robusto.

**5. Conclusão:**

Para maximizar a disponibilidade de equipamentos como prensas de papel pulp, os modelos GRU são as melhores opções. Eles não só proporcionam uma melhor precisão preditiva, mas também são mais eficientes em termos de tempo de treinamento e complexidade computacional.

### Discussão

A utilidade das redes neurais profundas na previsão de séries temporais é ressaltada pelo rápido desenvolvimento de novos modelos e o aprimoramento do poder computacional. Entre esses modelos, Long Short-Term Memory (LSTM) e Gated Recurrent Unit (GRU) destacam-se como escolhas principais para tarefas de previsão. Embora arquiteturas tradicionais ainda dominem grande parte do estado da arte, o apelo dos modelos LSTM e GRU reside na sua capacidade de lidar eficazmente com dados sequenciais.

**GRU vs. LSTM:**

As redes GRU oferecem várias vantagens sobre os modelos LSTM. Elas possuem uma arquitetura mais simples, o que as torna mais fáceis de treinar e menos intensivas em termos de computação. As GRUs também suportam taxas de reamostragem mais altas e desempenham bem tanto em conjuntos de dados pequenos quanto grandes. Esta pesquisa destaca que as GRUs, com suas camadas de codificação e decodificação, fornecem previsões precisas para aplicações industriais, como prever o comportamento de uma prensa de papel industrial 30 dias à frente com um Erro Percentual Absoluto Médio (MAPE) geralmente abaixo de 10%.

**Otimização do Modelo:**

A otimização dos modelos GRU mostrou gerar os melhores resultados com configurações específicas: uma janela deslizante de 12 dias, um período de amostragem de 1 hora e 50 unidades na camada oculta. A escolha das funções de ativação também desempenha um papel crucial, com a combinação ReLU-tanh emergindo como uma das mais eficazes. O estudo revela que treinar um único modelo para prever múltiplas variáveis de saída é mais vantajoso do que otimizar modelos individuais para cada variável. Essa abordagem holística aproveita as interdependências entre as variáveis, melhorando a precisão preditiva geral.

**Desempenho Comparativo:**

O modelo GRU demonstra desempenho superior ao modelo LSTM neste estudo. Ele não só atinge uma melhor precisão preditiva, mas também exibe robustez contra problemas comuns como gradientes explosivos ou desaparecidos. Essas características permitem que o modelo GRU aprenda eficazmente em uma gama mais ampla de configurações. Estudos anteriores mostraram que os modelos GRU são preditores superiores em contextos de dados univariados. Este estudo estende essas descobertas para dados multivariados, mostrando a eficácia do GRU em lidar com dados complexos e multivariados de séries temporais. Enquanto modelos como ARIMA podem ter um bom desempenho em pequenos conjuntos de dados com horizontes de previsão curtos, eles falham ao lidar com previsões mais longas e conjuntos de dados maiores, onde os modelos GRU se destacam.

### Conclusões

No setor industrial, minimizar o tempo de inatividade é crucial para manter a produtividade e reduzir perdas. A manutenção preditiva, facilitada por modelos avançados de IA, é fundamental para prever e mitigar possíveis falhas de equipamentos. Este estudo destaca as vantagens do uso de modelos GRU sobre modelos LSTM para prever os valores futuros dos dados dos sensores de uma prensa de papel industrial.

Os resultados sugerem que os modelos GRU não apenas requerem menos dados, mas também oferecem melhor desempenho preditivo com uma gama mais ampla de configurações de parâmetros. Pesquisas futuras se concentrarão em explorar o desempenho dos modelos GRU com diferentes intervalos de tempo para determinar configurações ótimas para diferentes horizontes de previsão.

Ao aproveitar as forças das redes GRU, as indústrias podem aprimorar suas estratégias de manutenção preditiva, levando a uma redução do tempo de inatividade e a uma maior eficiência operacional.