

SÍNTESE DE TRABALHO SOBRE A MODELAGEM DE SÉRIES TEMPORAIS DE TEMPERATURA COM APLICAÇÃO DE DEEP LEARNING - LSTM: UM ESTUDO DE CASO EM RIVERA, URUGUAI

LETICIA BARROS DIAS SOARES¹; PAMELA GABRIELA B. PEREIRA²;
GEILSON DE ALMEIDA SOARES³; SOFIA PAGLIARINI⁴; MARLON MAURICIO
HERNANDEZ CELY⁵; JAIRO VALÕES DE ALENCAR RAMALHO⁶.

¹Universidade Federal de Pelotas – LeticiaBarros1996@yahoo.com.br

²Universidade Federal de Pelotas – pamela.barboza@utec.edu.uy

³Universidade Federal de Pelotas – geilson.soares@ufpel.edu.br

⁴Universidade Federal de Pelotas – sofiapagliarini@gmail.com

⁵Universidade Federal de Pelotas – marlon.cely@ufpel.edu.br

⁶Universidade Federal de Pelotas – jairo.ramalho@ufpel.edu.br

1. INTRODUÇÃO

A energia geotérmica superficial é uma fonte de energia renovável promissora, oferecendo soluções sustentáveis para o aquecimento e resfriamento de edificações. Por exemplo, o sistema de Trocador de Calor Solo-Ar (TCSA) aproveita essa energia ao utilizar ventiladores de baixa potência que circulam ar por dutos subterrâneos, contribuindo para a redução do consumo de energia dos sistemas convencionais de ar-condicionado (LIU et al., 2023). A eficiência desse sistema depende de dados sobre as temperaturas do ar e do solo, que podem ser obtidos através de modelagem matemática e técnicas de inteligência artificial.

A previsão de séries temporais é uma área importante na análise de dados, com aplicações que vão desde a economia até a meteorologia. Séries temporais consistem em sequências de dados coletados ao longo do tempo, permitindo que analistas identifiquem padrões e tendências para prever eventos futuros. Nos últimos anos, redes neurais têm emergido como uma potente ferramenta para melhorar soluções específicas nesse campo (RIZVI, 2024).

Redes neurais são modelos computacionais inspirados no funcionamento do cérebro humano, capazes de aprender padrões complexos em grandes volumes de dados. ZHENG (2023) e GRAVES (2012) salientam que, com o advento da aprendizagem profunda, modelos como Redes Neurais Recorrentes (RNN) e *Long Short-Term Memory* (LSTM) se destacaram na tarefa de modelar dependências temporais em dados sequenciais, oferecendo resultados superiores em comparação com métodos tradicionais.

Este trabalho tem como objetivo prever o comportamento da temperatura do ar e do solo na cidade de Rivera (Uruguai). Tais temperaturas foram coletadas através de um sistema de monitoramento remoto, no período de julho de 2023 até abril de 2024, em parceria com a Universidad Tecnológica del Uruguay (UTEC).

2. METODOLOGIA

2.1 Coleta de dados

Este trabalho teve como objetivo prever a temperatura utilizando modelos de séries temporais com regressão linear e redes neurais recorrentes. Para isso, os dados de temperatura foram coletados por sensores instalados próximos à superfície do solo e nas profundidades de 0,4; 0,8; 1,2; 1,6 e 2 metros. Como o sistema de monitoramento requer manutenção periódica, pode haver períodos em

que os dados não estejam disponíveis. Os dados foram organizados por hora cumulativa, resultando em uma base de dados composta por 6056 registros, abrangendo um período contínuo de dez meses. Além disso, os dados passaram por pré-tratamento para remover valores nulos e inconsistentes, garantindo a qualidade das informações.

2.2 Escolha do modelo

Para a construção dos modelos preditivos, optou-se por empregar o método de Regressão Linear com aplicação de médias móveis, foi adotada para explorar uma abordagem mais simples e direta para a previsão, aproveitando a linearidade observada em partes do conjunto de dados, onde foram acumuladas 1816 horas. Além disso, foi utilizada uma rede neural recorrente do tipo LSTM (*Long Short-Term Memory*), devido à sua capacidade de capturar padrões de longo e curto prazo em séries temporais complexas.

2.3 Programação

Para assegurar a instalação adequada do ambiente de desenvolvimento integrado (IDE) *Spyder* e da linguagem de programação *Python*, foi utilizado o software *Anaconda Navigator*. As ferramentas e bibliotecas utilizadas incluem os pacotes *Numpy*¹, *Pandas*², *Keras*³ e *Scikit-Learn*⁴ para análise de dados, que oferecem estruturas de armazenamento eficientes e métodos avançados para a manipulação de dados.

2.4 Desenvolvimento da Rede

Para realizar as previsões de temperatura, foram utilizados dois modelos preditivos: Regressão Linear e uma Rede Neural Recorrente com técnica do LSTM. O modelo Regressão, foi empregada a média móvel de 24 horas como variável preditora, permitindo capturar variações locais de curto prazo. Para suavização e preenchimento de lacunas, foi utilizada a interpolação polinomial de ordem 12. Os dados foram divididos em conjuntos de treino e teste, utilizando as últimas 168 horas (equivalente há 7 dias) de observações para validação do desempenho dos modelos preditivos, enquanto o restante foi reservado para o treinamento. O conjunto de treino continha 70% dos dados, totalizando 4239 horas. As últimas 1816 horas foram previstas. Durante o treinamento, a média móvel foi calculada ao longo de uma janela de 24 horas, proporcionando uma visão mais equilibrada das variações diárias.

Para a Rede Neural Recorrente - LSTM Os dados de temperatura foram inicialmente normalizados para escalar os valores entre 0 e 1, o que ajuda a melhorar o desempenho da rede neural. Sequências de 24 horas (equivalentes a um dia) foram criadas para treinar o modelo, permitindo que ele aprendesse padrões sazonais e de variação diária. Essas sequências foram usadas como entradas (x), enquanto o valor da temperatura na próxima hora foi usado como saída (y). A estrutura do modelo LSTM consistiu em uma camada de entrada com 100 unidades LSTM, configurada para retornar sequências, seguida por um

¹ <https://numpy.org>

² <https://pandas.pydata.org>

³ <https://keras.io>

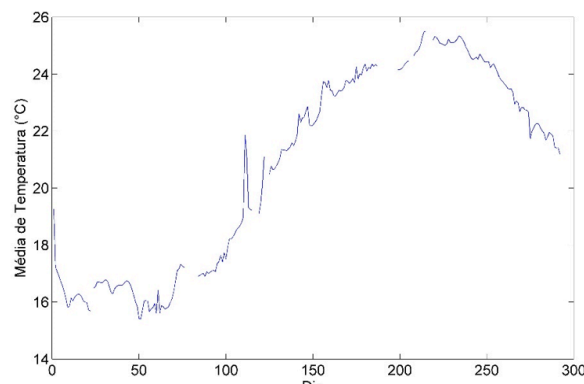
⁴ <https://scikit-learn.org/stable>

*Dropout*⁵ de 20% para prevenir o *Overfitting*⁶. Finalmente, uma camada densa com uma unidade foi usada para gerar a previsão final. O modelo foi compilado utilizando o otimizador 'Adam'⁷ e a função de perda 'Erro Médio Quadrático'. Para validação, os dados foram divididos em 80% para treino e 30% para teste. O modelo foi treinado por 50 épocas com um *batch size*⁸ de 32, permitindo ajustes iterativos nos pesos da rede para minimizar o erro. Após o treinamento, as previsões foram feitas sobre o conjunto de teste, e os resultados foram desnormalizados para a escala original. A avaliação do desempenho foi feita através de métricas de erro, como o Erro Absoluto Médio (MAE) e a Raiz do Erro Quadrático Médio (RMSE).

3. RESULTADOS

A título de ilustração, a Figura 1 mostra as médias de temperatura diárias no solo da UTEC, em uma profundidade de 1,2 m. Essa foi a profundidade tomada para análise. Como se pode perceber, há sete períodos onde as temperaturas estão faltando. Isso ocorreu devido a falhas pontuais nos sensores, onde ocorreram. Os resultados ficaram dentro do esperado para o local, com a temperatura do solo variando entre 15 e 26°C. A Figura 2 ilustra os gráficos de previsão da Regressão Linear e do LSTM para as temperaturas nas últimas 1816 horas de medição (correspondendo aos 30% de previsão).

Figura 1: Gráfico das médias diárias de temperatura do solo na UTEC a 1,2m de profundidade



Fonte: Dados do autor

A partir dos resultados apresentados na Tabela 1, observou-se que os dois métodos apresentaram Erros Médios Absolutos inferiores a 0,2 °C. Considerando que a temperatura média é próxima de 20 °C, isso representa erros relativos abaixo de 1%. Portanto, todos os métodos demonstraram precisão na simulação do fenômeno. Além disso, a análise revelou que a regressão linear teve um desempenho ligeiramente superior em relação aos dados observados, em termos de erros.

⁵ Desativação temporária de uma determinada fração de neurônios da rede.

⁶ *Overfitting* ocorre quando um algoritmo se adapta excessivamente aos dados de treinamento.

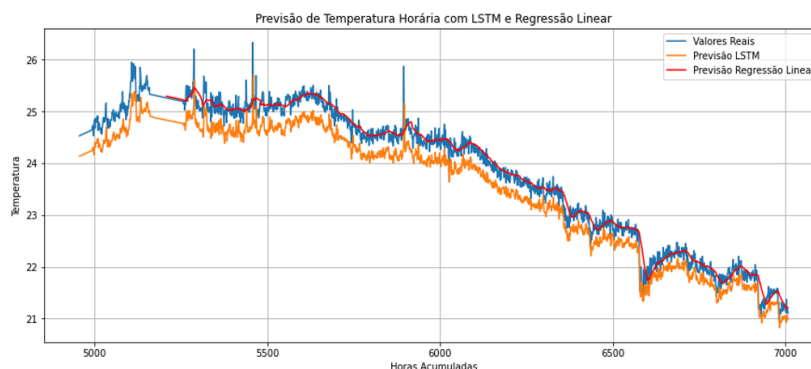
⁷ Ajusta a taxa de aprendizado de forma adaptativa para cada parâmetro.

⁸ *Batch Size* é a quantidade de dados que são processados em cada iteração de um algoritmo.

Tabela 1: Avaliação do desempenho do modelo

Modelo	Erro médio absoluto	Raiz do erro quadrático médio
Regressão Linear	0,111 °C	0,154 °C
RNN (LSTM)	0,169 °C	0,202 °C

Figura 2: Previsão de temperatura com LSTM e Regressão Linear



Fonte: Dados do autor

4. CONCLUSÕES

Apesar dos bons resultados, reforçam que a pesquisa está contribuindo para o desenvolvimento de modelos mais eficientes na detecção de mudanças nos padrões de temperatura. Futuramente, os modelos serão testados técnicas inovadoras como transformers, com períodos de dados mais longos para validar a robustez das previsões, proporcionando um melhor entendimento das variações de temperatura e ampliando suas aplicações práticas.

5. REFERÊNCIAS BIBLIOGRÁFICAS

Amaral, F. (2023). **Séries Temporais e Análises Preditivas com Python**. Disponível em: www.udemy.com. Acesso em: 24 jul. 2024.

GRAVES, A. **Long Short-Term Memory**. In: Rotulagem de sequência supervisionada com redes neurais recorrentes. Estudos em Inteligência Computacional, v. 385. Berlim, Heidelberg: Springer, 2012. Disponível em: https://doi.org/10.1007/978-3-642-24797-2_4. Acesso em: 23 jul. 2024.

RIZVI, Mohd Faizan; SAHU, Shivang; RANA, Sadhana. **ARIMA model time series forecasting**. International Journal for Research in Applied Science & Engineering Technology (IJRASET), v. 12, n. V, p. 3782, maio 2024. Disponível em: <https://www.ijraset.com>. Acesso em: 23/09/2024.

LIU, Z.; XIE, M.; ZHOU, Y.; HE, Y.; ZHANG, L.; ZHANG, G.; CHEN, D. **A state-of-the-art review on shallow geothermal ventilation systems with thermal performance enhancement system classifications, advanced technologies and applications**. Energy and Built Environment, v.4, n.2, p.148–168, 2023.