

LSTM e Transformers para Previsão de Temperatura: Um Estudo de Caso em Rivera, Uruguai

Roney M. Jr. Meirelles¹, Jairo V. de. A. Ramalho², Leticia B. D. Soares³, João I. M. Bezerra⁴, Marlon M. H. Cely⁵, Geilson de A. Soares⁶

UFPEL/Universidade Federal de Pelotas, Pelotas, RS

Oscar M. H. Rodriguez⁷, Carlos M. R. Diaz⁸

USP/Universidade de São Paulo, São Carlos, SP

Honório J. Fernando⁹

Universidade Federal Fluminense ICEX-UFF, Volta Redonda, RJ

Pamela G. B. Pereira¹⁰, Nibia R. N. Pastor¹¹, Marcos R. B. Olivera¹², Wilmar A. P. Castiblanco¹³, Cindy J. O. Gamba¹⁴

Universidad Tecnológica del Uruguay (UTEC), Rivera.

Resumo. A energia geotérmica superficial é uma fonte promissora de energia renovável, oferecendo soluções sustentáveis para o aquecimento e o resfriamento de edificações. Os trocadores de calor solo-ar (TCSA) exemplificam essa aplicação ao conectar sistemas de ventilação a dutos subterrâneos, reduzindo o consumo energético com climatização. No entanto, a avaliação de sua eficiência requer a análise de séries temporais das temperaturas do ar e do solo, que podem ser modeladas por técnicas matemáticas e inteligência artificial. Com os avanços da aprendizagem profunda, modelos como Long Short-Term Memory (LSTM) e Transformers têm se destacado na modelagem e previsão de dependências temporais. Este estudo investiga o comportamento das temperaturas do ar e do solo em Rivera, Uruguai, utilizando dados coletados entre julho de 2023 e abril de 2024 por um sistema de monitoramento remoto instalado na Universidad Tecnológica del Uruguay (UTEC). Vale ressaltar que o modelo LSTM se destacou nos resultados obtidos, apresentando uma Raiz do Erro Médio Quadrático (RMSE) de 0,27, o que indica alta capacidade preditiva e menor erro nas previsões.

Palavras-chave. Trocadores de Calor Solo-Ar, Inteligência Artificial, LSTM, Transformers, Previsão de Temperaturas

1 Introdução

A transição para fontes de energia renovável, como a geotérmica, hidrelétrica e biomassa, tem se intensificado, oferecendo alternativas viáveis às energias fósseis, responsáveis pelas emissões de gases

¹rjmeirelles999@gmail.com

²j.v.a.ramalho@gmail.com

³leticiabarros1996@yahoo.com.br

⁴jimbezerra@inf.ufpel.edu.br

⁵marlon.cely@ufpel.edu.br

⁶geilsonsrz@gmail.com

⁷oscarahr@sc.usp.br

⁸carlosruiz978@usp.br

⁹honoriofernando@id.uff.br

¹⁰pamela.barboza@utec.edu.uy

¹¹nibia.navarro@utec.edu.uy

¹²marcos.burlon@utec.edu.uy

¹³wilmar.pineda@utec.edu.uy

¹⁴cindy.ortiz@utec.edu.uy

de efeito estufa [11]. Dentro desse contexto, os trocadores de calor solo-ar (TCSA) se destacam por utilizarem energia geotérmica renovável e ventiladores de baixa potência para conectar sistemas de ventilação a dutos enterrados, reduzindo a carga térmica de edificações [14] e, assim, diminuindo o consumo energético em comparação aos sistemas tradicionais de climatização, o que contribui para a sustentabilidade e eficiência energética.

A estimativa da eficiência dos TCSA depende da modelagem da interação térmica entre o ar e o solo, assim como da obtenção de dados experimentais. No entanto, falhas no sensoriamento podem comprometer a coleta dessas informações [5], [9]. Para este estudo foram levados em consideração somente dados completos de um intervalo da série temporal estudada, para usá-los nos modelos preditivos Long Short-Term Memory (LSTM) e Transformers, garantindo maior continuidade e confiabilidade na análise da série temporal. Essas técnicas têm demonstrado avanços significativos na modelagem de dependências temporais e previsão de dados [3].

A arquitetura Transformer, aplicada no trabalho [1], tem sido útil no desenvolvimento de modelos de linguagem avançados como o GPT-3. Esse modelo se destaca pelo grande volume de dados e parâmetros, o que amplia sua capacidade de generalização e adaptação a novas tarefas, muitas vezes sem ajustes complexos. Essa abordagem tem influenciado o desenvolvimento de modelos de aprendizado profundo para predição. A aplicação do Long Short-Term Memory (LSTM) também tem crescido nos últimos anos, especialmente em áreas como a indústria de energia [8] e a bio-médica. O uso de LSTM em séries temporais tem se mostrado eficaz na previsão e detecção de padrões complexos.

Neste contexto, este trabalho tem como objetivo prever o comportamento das temperaturas do ar e do solo na cidade de Rivera, Uruguai, utilizando modelos LSTM e Transformers. Os dados de temperatura foram coletados por sensores em diferentes profundidades ao longo de dez meses, nas dependências da Universidad Tecnológica del Uruguay (UTEC). Em particular, o trabalho apresenta resultados referentes aos sensores localizados a 0,4 m de profundidade.

2 Referencial Teórico

2.1 Sistema de Trocador de Calor Solo-Ar

Um esquema dos TCSA em um dia frio é ilustrado na Figura 1. Aqui, as cores azul e vermelho representam temperaturas frias e quentes respectivamente. Na pesquisa sobre TCSA, determinar o potencial térmico do solo, ou seja, a diferença entre a temperatura do solo e a do ar, é importante para avaliar a viabilidade de construção do sistema. Este trabalho visa contribuir com a modelagem desse potencial na cidade de Rivera, a partir da análise de dados de sensoriamento remoto local.

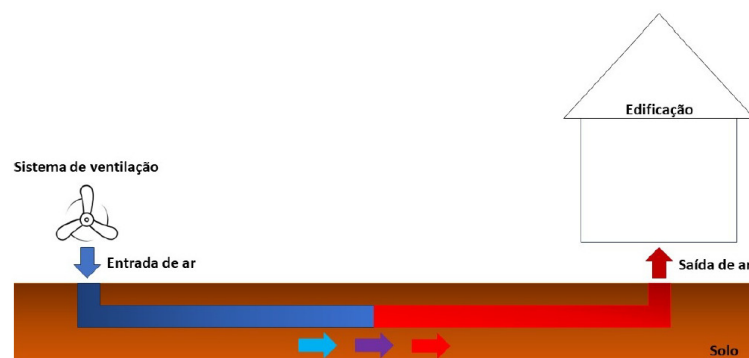


Figura 1: Esquema de funcionamento dos TCSA. Fonte: Próprio autor, 2024.

2.2 Séries Temporais

Séries temporais são sequências de dados coletados ao longo do tempo, nas quais a ordem das observações é relevante para a análise. Comumente usadas em áreas como economia, meteorologia e aprendizado de máquina, elas visam modelar padrões e prever tendências. Usualmente, as séries temporais têm uma natureza dinâmica e não estacionária, isto é, as médias ou variâncias dos dados variam ao longo do tempo, demandando técnicas específicas para sua modelagem (ver, por exemplo [7]). Dito isso, este estudo envolve a utilização dos modelos Transformers e LSTM aplicados em uma série temporal $T(t_1), T(t_2), \dots, T(t_n)$, onde as observações de temperaturas T são realizadas de hora em hora, em instantes de tempo t_1, t_2, \dots, t_n . A série é composta por valores medidos na UTEC de Rivera.

2.3 Transformers

Os Transformers são uma técnica de aprendizado profundo baseada em mecanismos de atenção [12], permitindo que modelos processem sequências de dados de forma paralela e eficiente. Esse modelo utiliza a atenção autoregressiva para capturar dependências de longo alcance nos dados, além de ter um grande impacto no processamento de linguagem natural (PLN), sendo amplamente adotada em tarefas como tradução automática e geração de texto. Sua flexibilidade e escalabilidade fazem com que seja uma técnica aplicável a diversos domínios além do PLN.

Em séries temporais, Transformers modelam padrões complexos e capturam relações de longo alcance. O mecanismo de atenção permite ao modelo focar nas partes mais relevantes da entrada para cada previsão. A atenção é calculada com a função *softmax* para normalizar os pesos, de modo que sua soma seja igual a 1. O *softmax* é uma função matemática que converte os valores em uma distribuição de probabilidade, garantindo que todos os pesos somem para 1. A equação para calcular a atenção é dada por:

$$\text{Atenção}(Q, C, V) = \text{softmax} \left(\frac{QC^T}{\sqrt{d_c}} \right) V. \quad (1)$$

Onde:

- Q é a matriz de consultas,
- C é a matriz de chaves,
- V é a matriz de valores,
- d_c é a dimensionalidade das chaves.

A expressão $\frac{QC^T}{\sqrt{d_c}}$ representa a similaridade entre a consulta (Q) e as chaves (C), e a função *softmax* é aplicada para normalizar essa similaridade. O *softmax* transforma essas similaridades em uma distribuição de pesos, onde cada valor é um número entre 0 e 1, e a soma de todos os pesos resulta em 1. Esses pesos são então aplicados aos valores (V) para gerar a saída ponderada, permitindo que o modelo se concentre nas partes mais relevantes da entrada.

Essa capacidade de focar em diferentes partes da entrada, dependendo do contexto, torna os Transformers uma alternativa eficaz para aplicações como previsão de demanda, variação climática e análise financeira. Um dos primeiros trabalhos a explorar Transformers para séries temporais foi o estudo de [2], onde foi introduzido o modelo *Temporal Fusion Transformer* (TFT), que demonstrou melhorias em relação a métodos tradicionais. Este artigo propõe a aplicação do modelo Transformers para previsão de séries temporais, considerando que é uma abordagem de aprendizado profundo amplamente utilizada em problemas de detecção e previsão [4].

2.4 Long Short-Term Memory (LSTM)

As redes LSTM foram desenvolvidas por [6] como uma alternativa para superar as limitações das redes neurais recorrentes (RNNs) convencionais, especialmente os desafios relacionados ao desaparecimento e à explosão do gradiente durante o processo de retropropagação no tempo [13]. Sua estrutura permite que essas redes armazenem e utilizem informações por períodos prolongados, o que as torna ideais para a previsão de séries temporais e a modelagem de sistemas dinâmicos. O funcionamento das LSTM [10] é baseado em três portas: a porta de entrada i_t , a porta de esquecimento f_t e a porta de saída o_t , que controlam o fluxo de informações através da célula de memória C_t e a saída h_t . As atualizações da célula de memória e da saída são descritas pelas seguintes equações:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (2)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \quad (4)$$

$$h_t = o_t \cdot \tanh(C_t). \quad (5)$$

Onde σ é a função sigmoide e \tanh é a função tangente hiperbólica. A função de esquecimento f_t decide quais informações da célula de memória anterior C_{t-1} devem ser mantidas, a função de entrada i_t controla a quantidade de nova informação a ser armazenada na célula, e a função de saída o_t regula a informação que será passada para a próxima camada ou saída. Essas operações matemáticas permitem que as redes LSTM armazenem dependências de longo prazo e evitem os problemas de desaparecimento do gradiente encontrados em redes neurais recorrentes tradicionais.

3 Metodologia

3.1 Coleta de Dados

As medições foram feitas por conjuntos de sensores instalados a diferentes profundidades. Em particular, este trabalho foca em uma amostra da pesquisa, relativa a três dispositivos enterrados a 0,4 metros de profundidade. O uso de múltiplos dispositivos teve o objetivo de garantir redundância e confiabilidade nos dados. A base de dados composta por 6.056 registros ao longo de dez meses, foi pré-processada para remover valores nulos e inconsistentes. Esses valores inconsistentes são dados que apresentavam erros de temperatura causados por falhas nos sensores que indicava temperaturas de verão durante o inverno, logo, tais leituras foram excluídas da base, garantindo a qualidade das previsões. Foram calculadas médias horárias das medições e comparados os valores registrados. Quando a diferença entre os dispositivos não ultrapassava 1°C, a média de todos os três era calculada. Se um dos dispositivos apresentava um valor discrepante, a média era calculada com base em dois. Quando os três apresentavam valores díspares, a medição era considerada inválida.

3.2 Programação

As implementações foram realizadas em Python, utilizando bibliotecas como NumPy e Pandas, que permitem a organização e manipulação de grandes volumes de dados. Para a construção e treinamento dos modelos de redes neurais, foram utilizadas as bibliotecas Keras e Scikit-Learn. A integração dessas ferramentas permitiu a implementação dos algoritmos de previsão. Os dados foram então segmentados em conjuntos de treinamento e teste, 83.33% e 16.76 % respectivamente para realizar a previsão das temperaturas no intervalo entre os pontos 5.000 e 6.000. A avaliação dos modelos foi realizada por meio de métricas como Erro Absoluto Médio (MAE), Erro Médio

Absoluto (MSE), Raiz do Erro Médio Quadrático (RMSE) e Coeficiente de Determinação (R^2), permitindo uma análise do desempenho preditivo dos modelos aplicados no estudo.

As métricas de avaliação utilizadas, como MAE, MSE, RMSE e R^2 , são definidas pelas seguintes fórmulas:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (6)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (8)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (9)$$

onde y_i são os valores reais, \hat{y}_i são os valores previstos, \bar{y} é a média dos valores reais e n é o número total de amostras.

4 Resultados

A Tabela 1 apresenta as métricas de validação dos modelos LSTM e Transformers para a previsão das temperaturas a 0,4 metros de profundidade. O modelo Transformers obteve melhor desempenho no MAE (0.20) e R^2 (0.91), indicando maior precisão nas previsões e melhor capacidade de explicação da variabilidade dos dados. Por outro lado, o modelo LSTM teve desempenho superior no MSE (0.07) e RMSE (0.27), sugerindo que foi mais robusto a grandes desvios.

Tabela 1: Resultados das métricas para cada um dos modelos.

Modelos	MAE	MSE	RMSE	R^2
LSTM	0.25	0.07	0.27	0.80
Transformers	0.20	0.23	0.48	0.91

Na Tabela 1, o modelo LSTM alcançou o menor valor da métrica Raiz do Erro Médio Quadrático (RMSE), o que implica que foi o melhor modelo para previsão no estudo em comparação com o modelo Transformers, uma vez que essa métrica indica que o modelo tem alta capacidade preditiva e possui um baixo erro de previsão. Na Figura 2, mostra a previsão dos modelos LSTM e Transformers, no qual o LSTM se destacou como melhor modelo para previsão desse estudo.

Para validar os resultados de previsão no intervalo entre 5.000 e 6.000, analisou-se o gráfico da Figura 2, que compara as estimativas dos modelos LSTM (preto) e Transformers (laranja) com os dados reais (azul). O desempenho é avaliado pela proximidade entre previsões e valores observados. Observa-se que o LSTM apresentou maior aderência aos dados reais, indicando melhor desempenho no intervalo considerado. Conclui-se, portanto, que ambos os modelos são eficazes na previsão de séries temporais, reproduzindo com precisão os padrões da temperatura dos trocadores de calor solo-ar analisados.

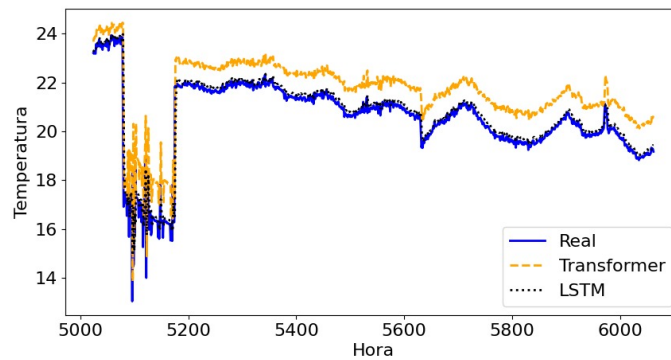


Figura 2: Previsão das temperatura com os modelos LSTM e Transformers. Fonte: Próprio autor, 2025.

5 Considerações Finais

Este trabalho utilizou os modelos LSTM e Transformers para a previsão de temperatura. Os resultados demonstraram que ambos os modelos foram eficazes na previsão dos dados, com desempenho consistente em relação aos valores reais. No entanto, o modelo LSTM apresentou melhor desempenho do que o modelo Transformers.

A aplicação de IA em séries temporais mostrou-se eficaz no estudo dos Trocadores de Calor Solo-Ar. Os resultados destacam a eficiência dos modelos e a importância de integrar dados meteorológicos e métodos preditivos para aprimorar a análise dos sistemas. Este trabalho contribui para literatura estudada ao evidenciar o potencial dessas abordagens em aplicações prática.

Como sugestão para trabalhos futuros, propõe-se o estudo da aplicação de modelos de inteligência artificial para preencher os valores nulos que foram removidos da base de dados utilizada neste estudo de previsão. O objetivo é imputar esses valores ausentes por meio de técnicas de IA, possibilitando a comparação dos resultados de previsão obtidos com os dados completados em relação aos resultados alcançados neste trabalho.

Agradecimentos

Roney Meirelles, João Bezerra, Marlon Cely, Oscar Rodriguez e Carlos Diaz agradecem a bolsa da Petrobras (LEMI- Petrobras). Leticia Soares agradece a UFPel por sua bolsa de iniciação científica. Jairo Ramalho agradece a Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) pela concessão de apoio financeiro via EDITAL FAPERGS 09/2023, PROGRAMA PESQUISADOR GAÚCHO – PqG, termo de outorga 24/2551-0001485-3. Todos os autores agradecem o apoio da UTEC de Rivera pelo apoio a construção do experimento.

Referências

- [1] T. B. Brown et al. “Language models are few-shot learners”. Em: **Proceedings of the 34th International Conference on Neural Information Processing Systems**. NIPS '20. Vancouver, BC, Canada: Curran Associates Inc., 2020. ISBN: 9781713829546.
- [2] L. Bryan et al. “Temporal Fusion Transformers for interpretable multi-horizon time series forecasting”. Em: **Machine Learning** 3 (2020), p. 27. DOI: 10.1007/s40314-014-0163-6.

- [3] F. Chollet. **Deep Learning with Python**. Second Edition. Manning Publications, 2021. ISBN: 978-1617296864.
- [4] N. M. Foumani et al. “Improving position encoding of transformers for multivariate time series classification”. Em: **Data Mining and Knowledge Discovery** 38 (2024), pp. 22–48. DOI: 10.1007/s10618-023-00948-2.
- [5] A. Gaddam, T. Wilkin e M. Angelova. “Modelos de detecção de anomalias para detecção de falhas de sensores e outliers na IoT - Uma pesquisa”. Em: **13ª Conferência Internacional sobre Tecnologia de Sensoriamento (ICST)**. 2019, pp. 1–6. DOI: 10.1109/ICST46873.2019.9047684.
- [6] S. Hochreiter e J. Schmidhuber. “Long short-term memory”. Em: **Neural Computation** 9.8 (1997), pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.
- [7] G. Kirchgaessner, R. Thiele e M. J. Weidmann. **Introduction to Modern Time Series Analysis**. Berlin, Heidelberg: Springer, 2009. ISBN: 978-3-540-49059-9.
- [8] G. N. Kouziokas. “Long Short-Term Memory (LSTM) Deep Neural Networks in Energy Appliances Prediction”. Em: **Conferência Pan-Helênica sobre Eletrônica e Telecomunicações (PACET)**. 2019, pp. 1–5. DOI: 10.1109/PACET48583.2019.8956252.
- [9] X. Li et al. “Remaining Useful Life Prediction With Partial Sensor Malfunctions Using Deep Adversarial Networks”. Em: **IEEE/CAA Journal of Automatica Sinica** 10.1 (2023), pp. 121–134. DOI: 10.1109/JAS.2022.105935.
- [10] I. Malashin et al. “Applications of long short-term memory (LSTM) networks in polymeric sciences: A review”. Em: **Polymers** 16.18 (2024). DOI: 10.3390/polym16182607.
- [11] M. F. Rabbi et al. “Energy security and energy transition to achieve carbon neutrality”. Em: **Energies** 15.21 (2022). DOI: 10.3390/en15218126.
- [12] A. Vaswani et al. “Attention is all you need.” Em: **Adv. Neural Inf** 30 (2017), pp. 6000–6010.
- [13] P. J. Werbos. “Backpropagation through time: What it does and how to do it”. Em: **Proceedings of the IEEE** 78.10 (1990), pp. 1550–1560. DOI: 10.1109/5.58337.
- [14] H. Zeng, N. Diao e Z. Fang. “Efficiency of vertical geothermal heat exchangers in the ground source heat pump system”. Em: **Journal of Thermal Science** 12.1 (2003), pp. 77–81. DOI: 10.1007/s11630-003-0012-1.