

# Transformer

## Advantages:

- It improves precision and efficiency as it focuses on both local and global features [1].
- As a transformer has a temporal decoder, it solves long sequence tasks effectively and mitigates noise in time series [2].

## Disadvantages:

- It requires high computation and memory, which will be limited in the constrained resource settings [3].
- Over-parameterization and overfitting problems [4].
- It is difficult to train as it requires designing optimizers and learning rate schedulers carefully [5].

## Example of usage:

- **Code example**

Build your own Transformer from scratch using Pytorch [6]

- **Application example**

- Detecting the railway surface defect (RSDs) [1].
- The time-series transformer effectively deals with the classification problem with imbalanced temporal datasets [7].

## References

- [1] F. Guo, J. Liu, Y. Qian, and Q. Xie, "Rail surface defect detection using a transformer-based network," *Journal of Industrial Information Integration*, vol. 38, p. 100584, Feb. 2024, doi: <https://doi.org/10.1016/j.jii.2024.100584>.
- [2] C. Liu, S. He, H. Liu, J. Chen, and H. Dong, "WindTrans: Transformer-Based Wind Speed Forecasting Method for High-Speed Railway," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 6, pp. 4947–4963, Jun. 2024, doi: <https://doi.org/10.1109/tits.2023.3337150>.
- [3] Z. Luo, N. H. Yan, and N. X. Pan, "Optimizing Transformer Models for Resource-Constrained Environments: A Study on Model Compression Techniques," *Journal of Computational Methods in Engineering Applications*, vol. 3, no. 1, pp. 1–12, Nov. 2023, doi: <https://doi.org/10.62836/jcmea.v3i1.030107>.
- [4] J. Wu *et al.*, "UniDrop: A Simple yet Effective Technique to Improve Transformer without Extra Cost," *arXiv (Cornell University)*, Jan. 2021, doi: <https://doi.org/10.48550/arxiv.2104.04946>.
- [5] L. Liu, X. Liu, J. Gao, W. Chen, and J. Han, "Understanding the Difficulty of Training Transformers," *arxiv.org*, Apr. 2020, doi: <https://doi.org/10.48550/arXiv.2004.08249>.
- [6] A. Sarkar, "Build your own Transformer from scratch using Pytorch," *Medium*, Apr. 26, 2023.  
<https://medium.com/data-science/build-your-own-transformer-from-scratch-using-pytorch-h-84c850470dcb> (accessed Aug. 20, 2025).
- [7] Y. Liu, Sudanthi Wijewickrema, A. Li, C. Bester, S. O'Leary, and J. Bailey, "Time-Transformer: Integrating Local and Global Features for Better Time Series Generation," *Society for Industrial and Applied Mathematics eBooks*, pp. 325–333, Jan. 2024, doi: <https://doi.org/10.1137/1.9781611978032.37>.