

Estudio sobre la Efectividad del Positional Encoding en Transformers para Series Temporales y Diseño de Mecanismos Adaptados

Cristhian Moya Mota

Tutor: Julián Luengo Martín

Departamento de Ciencias de la Computación e Inteligencia Artificial

Cotutor: Diego Jesús García Gil

Departamento de Lenguajes y Sistemas Informáticos

8 de septiembre de 2025

Contenido

- 1 Introducción
- 2 Estado del arte
- 3 Propuestas de positional encoding
- 4 Evaluación de los PE mediante conjuntos de datos
- 5 Conclusiones y trabajos futuros
 - Conclusiones
 - Trabajos futuros

Contenido

- 1 Introducción
- 2 Estado del arte
- 3 Propuestas de positional encoding
- 4 Evaluación de los PE mediante conjuntos de datos
- 5 Conclusiones y trabajos futuros
 - Conclusiones
 - Trabajos futuros

Introducción

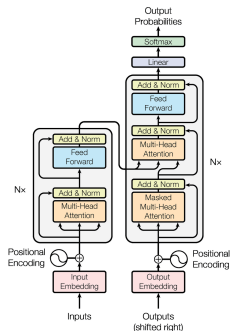
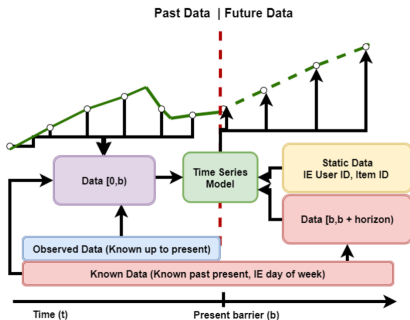


Figura: Forecasting en Series Temporales¹. Transformers²

¹<https://developer.nvidia.com/blog/time-series-forecasting-with-the-nvidia-time-series-prediction-platform-and-triton-inference-server/>

²<https://doi.org/10.48550/arXiv.1706.03762>

Introducción. Justificación

Aspectos que justifican la realización del proyecto:

- Falta de captura de la estructura. Ausencia de información semántica.
- Dificultad para adaptarse a diferentes escalas temporales y falta de semántica.
- Complejidad computacional y falta de interpretabilidad en la metodología.

Introducción. Objetivos

Este proyecto persigue:

- Comprender y exhibir las carencias de los métodos actuales.
- Proponer nuevos encodings para series temporales empleando Transformers.
- Evaluar la efectividad de las nuevas propuestas en diferentes ámbitos y conjuntos de datos.
- Identificar y establecer la base para nuevos métodos de codificación.

Contenido

1 Introducción

2 Estado del arte

3 Propuestas de positional encoding

4 Evaluación de los PE mediante conjuntos de datos

5 Conclusiones y trabajos futuros

- Conclusiones

- Trabajos futuros

Estado del arte. Arquitecturas y PE

- Métodos estadísticos: STL, ARIMA, Prophet.
- Métodos basados en Transformer:
 - Informer
 - Autoformer
 - FEDformer

Problemas:

- 1 Escasa adición de información local.
- 2 Mecanismos de atención poco cercanos a la semántica del dato.

¿Solución? → Crear una nueva familia de codificaciones posicionales, capaz de captar información local y global.

Estado del arte. Caso particular de Informer

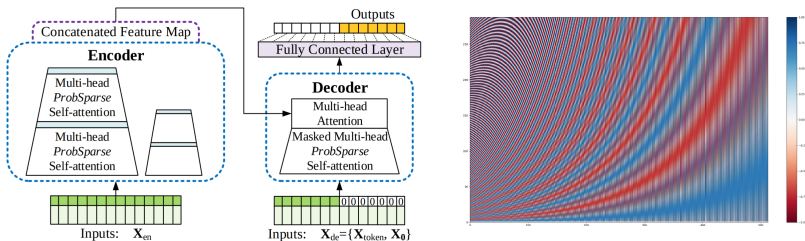


Figura: Informer: arquitectura y encoding empleado³

³<https://doi.org/10.48550/arXiv.2012.07436>

Contenido

1 Introducción

2 Estado del arte

3 Propuestas de positional encoding

4 Evaluación de los PE mediante conjuntos de datos

5 Conclusiones y trabajos futuros

- Conclusiones

- Trabajos futuros

Encodings propuestos: Aspectos clave

Para mejorar la calidad del encoding, hay dos alternativas:

- Modificar el mecanismo de atención → pérdida de información local y complejidad añadida
- Modificar únicamente el PE → permite aprovechar arquitecturas existentes y trabajar directamente sobre el dato

Resultado: familia de modelos WinStat:

- 1 Sólo modificar la codificación posicional
- 2 Concatenar información en lugar únicamente sumarla
- 3 Combinar PE existentes de manera ponderada (normalizada con Softmax):

$$X_{\text{final}} \leftarrow \tilde{w}_0 \cdot \tilde{X} + \tilde{w}_1 \cdot X_{pe1} + \tilde{w}_2 \cdot X_{pe2} + \tilde{w}_3 \cdot X_{pe3}$$

Encodings propuestos: WinStat (Variante base)

Ventana local de tamaño W fijo, en cada posición t :

- *Media*: $\mu_t = \frac{1}{|\mathcal{W}_t|} \sum_{x \in \mathcal{W}_t} x$
- *Desviación estándar*: $\sigma_t = \sqrt{\frac{1}{|\mathcal{W}_t|} \sum_{x \in \mathcal{W}_t} (x - \mu_t)^2}$
- *Mínimo*: $m_t^{\min} = \min_{x \in \mathcal{W}_t} x$
- *Máximo*: $m_t^{\max} = \max_{x \in \mathcal{W}_t} x$

$$s_t = [\mu_t, \sigma_t, m_t^{\min}, m_t^{\max}] \in \mathbb{R}^4$$

Y, finalmente, el embedding enriquecido de la posición t es:

$$\tilde{x}_t = [x_t \parallel s_t] \in \mathbb{R}^{d+4}$$

Encodings propuestos: WinStatLag

Parte de WinStat, añadiendo $|\mathcal{L}|$ retardos especificados:

$$s_t = [\mu_t, \sigma_t, m_t^{\text{mín}}, m_t^{\text{máx}}] \in \mathbb{R}^4$$

Definimos para cada $\ell_j \in \mathcal{L}$:

$$\delta_t^{(\ell_j)} = |x_t - x_{t-\ell_j}|, \quad \text{si } t - \ell_j \geq 1$$

El *embedding* enriquecido es la concatenación de s_t y $\delta_t^{(\ell_j)}$

$$\tilde{x}_t = [x_t \parallel s_t \parallel \delta_t^{(\ell_1)} \parallel \dots \parallel \delta_t^{(\ell_p)}] \in \mathbb{R}^{d+4+p}$$

Encodings propuestos: WinStatFlex

Basado en WinStatLags, añadiendo la información ponderada de:

- 1 Encoding sinusoidal original
- 2 LPE:

$$LPE_{(pos)} = W_{pos}, \quad W_{pos} \in \mathbb{R}^d$$

$$X_{LPE} = \tilde{x}_t + LPE[pos]$$

- 3 TAPE:

$$\omega_k^{new} = k \cdot \frac{d_{model}}{L} \quad (1)$$

$$\begin{aligned} \text{TAPE}_{(pos,2i)} &= \sin(pos \cdot \omega_i^{new}) \\ \text{TAPE}_{(pos,2i+1)} &= \cos(pos \cdot \omega_i^{new}) \end{aligned} \quad (2)$$



Encodings propuestos: WinStatTPE

- Emplea WinStatFlex como base
- Sustituye TAPE por TPE:

$$S(i, j) = \exp \left(-\frac{\|x_i - x_j\|^2}{2\sigma^2} \right)$$

Añadiéndose al encoding sinusoidal original la nueva componente:

$$T-PE(i) = PE(i) + S(i, j)$$

$$X_{TPE} = \tilde{x}_t + T - PE[pos]$$

Contenido

- 1 Introducción
- 2 Estado del arte
- 3 Propuestas de positional encoding
- 4 Evaluación de los PE mediante conjuntos de datos
- 5 Conclusiones y trabajos futuros
 - Conclusiones
 - Trabajos futuros

Evaluación. Conjuntos de datos

Conjuntos de datos para la evaluación:

- Household Power Consumption (HPC)
- ETTh1
- ETTh2
- Yellow Trip Data
- TINA

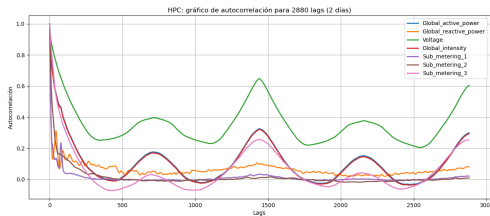


Figura: Autocorrelación de Household Power Consumption

Contenido

- 1 Introducción
- 2 Estado del arte
- 3 Propuestas de positional encoding
- 4 Evaluación de los PE mediante conjuntos de datos
- 5 Conclusiones y trabajos futuros
 - Conclusiones
 - Trabajos futuros

El proyecto ha conseguido:

Gracias



¿Preguntas?