

Gaussian Error Linear Units (GELUS)

📄 Integrantes:

- 📄 Henry Mantilla
- 📄 Alejandro Moreno

TLDR

GELU (Unidad Lineal de Error Gaussiano) es una función de activación definida como $x\Phi(x)$, donde $\Phi(x)$ es la función de distribución acumulada de una normal estándar. Puede interpretarse como una “puerta” suave que deja pasar x en proporción a qué tan probable es que sea mayor que una muestra de una normal estándar. A diferencia de ReLU, que corta en seco los valores negativos, GELU hace una transición suave entre bloquear valores negativos y dejar pasar los positivos, lo que mejora el flujo de gradientes. El artículo también propone aproximaciones rápidas usando tanh o una sigmoide escalada. Experimentalmente, GELU supera a ReLU y ELU en redes grandes (por ejemplo Transformers), logrando mejor convergencia y precisión con un costo computacional apenas mayor.

📄 Material de apoyo:

- 📄 **Diapositivas:** [Ver presentación](#)
- 📄 **Paper:** [Gaussian Error Linear Units \(GELUS\)](#)

Revision #4

Created 17 September 2025 20:13:36 by Henry

Updated 20 September 2025 16:18:43 by Henry