Uncertainty-aware Computation: From Error Propagation to Reliable Decision Support

Maria Lizeth Reyna Cruz
University of Texas at El Paso

Advisors: Martine Ceberio, Christoph Lauter

Workshop: LLM-Driven Code Generation and Automation for HPC and Scientific Computing

Outline

i. Error Analysis in Deep Neural Networks
ii. Uncertainty Modeling Approaches
iii. Uncertainty Applications in Medicine

Motivation

- LLMs are increasingly used in HPC and scientific workflows
- Code generation + automation + numerical pipelines
- But reliability is not only about correctness of code
- Floating-point errors still propagate in:
  - deep learning systems
  - scientific computation pipelines
  - reduced precision hardware

Problem

Even in correct neural networks:

- every operation introduces floating-point error
- errors accumulate across layers
- nonlinear functions amplify small perturbations

Key question: How do numerical errors propagate through deep networks?

Focus

We analyze error propagation in activation functions:

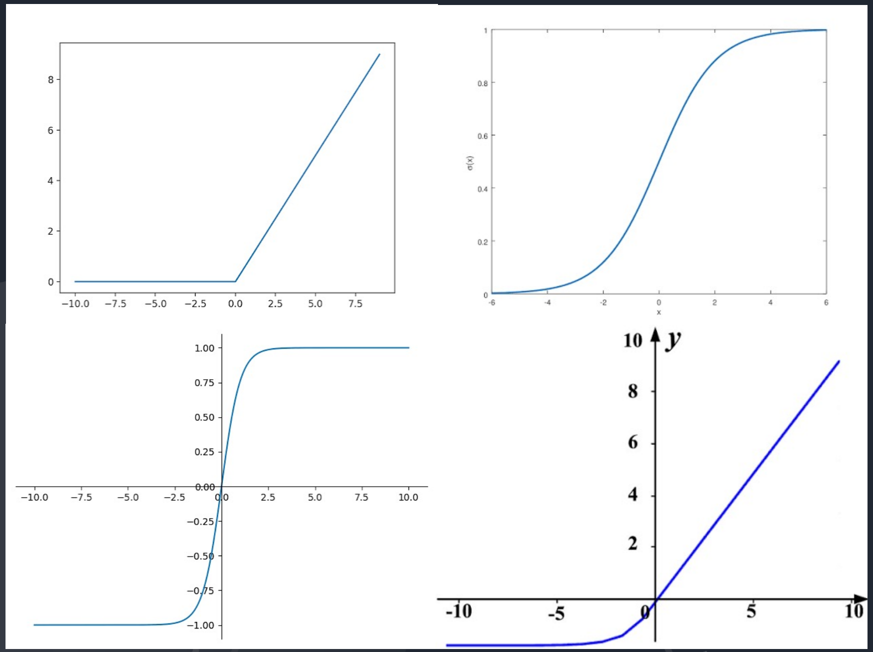
- ReLU
- sigmoid
- tanh
- ELU
- softmax

Using:

- absolute error
- relative error

Goal: quantify numerical sensitivity of nonlinearities

Preprint: Reyna-Cruz, M. L., Ruiz-Rohena, K., Guel, Y. I., Salgado, H., Taldir, L., Ramos, E. P., ... & Volkova, A. (2025). Contribution to Error Analysis of Deep Neural Networks: Case of the Activation Functions. Available at: https://inria.hal.science/hal-05367563/



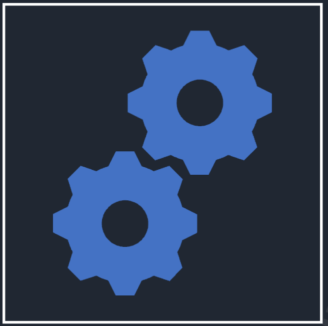
Activation functions behave like numerical operators

ReLU → piecewise linear → stable

sigmoid / tanh → nonlinear amplification

ELU → hybrid behavior

softmax → coupled vector instability



Activations are not just ML components
They define numerical stability of the system

Results

ReLU	stable under perturbations
	bounded error propagation
Sigmoid / tanh	nonlinear amplification
	sensitivity increases with input magnitude
ELU	intermediate stability
Softmax	interdependent error across dimensions

Activation	Absolute Error Bound	Relative Error Bound
ReLU $f(x) = \max(0, x)$	$\lvert\mathrm{ReLU}(x+\Delta x) - \mathrm{ReLU}(x)\rvert \leq \lvert\Delta x\rvert$. Exact piecewise characterization: if x and $x + \Delta x$ are in the same region, $\Delta y = \Delta x$ (if $x > 0$) or $\Delta y = 0$ (if $x \leq 0$).	For any perturbation δ, $$\lvert\mathrm{ReLU}(x+\delta) - \mathrm{ReLU}(x)\rvert \leq \lvert\delta\rvert.$$ When the relative ratio is defined, $$\frac{\lvert\Delta y\rvert}{\lvert\mathrm{ReLU}(x)\rvert} \leq \frac{\lvert\delta\rvert}{\lvert x\rvert} \quad (x > 0).$$ Undefined at $x = 0$.
Sigmoid $\sigma(x)$	Absolute bound: $$\lvert\sigma(x+\Delta x) - \sigma(x)\rvert \leq \frac{1}{4}\lvert\Delta x\rvert.$$ Sharper non-asymptotic estimate: $$\lvert\sigma(x+\Delta x) - \sigma(x)\rvert \leq \sigma(x)(1-\sigma(x))(e^{\lvert\Delta x\rvert} - 1).$$ Second-order refinement: $$\sigma(x+\Delta x) - \sigma(x) - \sigma'(x)\Delta x \leq \frac{1}{12\sqrt{3}}\lvert\Delta x\rvert^2.$$	If $\Delta x = x\,\varepsilon$ and $\lvert\varepsilon\rvert \leq \bar{\varepsilon} \leq \frac{1}{4}$, then $$\varepsilon' = (1-\sigma(x))\,x\,\varepsilon + O((x\varepsilon)^2),$$ and $$\lvert\varepsilon'\rvert \leq e^{\lvert x\rvert}\bar{\varepsilon} - 1.$$
Tanh $\tanh(x)$	$$\lvert\tanh(x+\Delta x) - \tanh(x)\rvert \leq (1-\tanh^2 x)\lvert\Delta x\rvert$$ $$+\lvert\tanh x\rvert(1-\tanh^2 x)\lvert\Delta x\rvert^2 + \frac{1}{3}(1-\tanh^2 x)\lvert 1$$ $$-3\tanh^2 x\rvert\,\lvert\Delta x\rvert^3 + 0.29535\,\lvert\Delta x\rvert^4.$$	If $\lvert\varepsilon\rvert \leq \bar{\varepsilon} \leq \frac{1}{4}$, then there exists $\bar{\varepsilon}' \leq 2.6262\,\bar{\varepsilon}$ such that $$\tanh(x(1+\varepsilon)) = \tanh(x)(1+\varepsilon') \quad \text{with} \quad \lvert\varepsilon'\rvert \leq \bar{\varepsilon}'.$$
ELU $\mathrm{elu}(x)$	Let $\alpha \in \mathbb{R}$ with $\alpha > 0$. Then $y\Delta$ is given by , $$y_\Delta = \begin{cases} 0, & x = \hat{x} \\ \lvert x\Delta\rvert, & x \geq 0\ \&\ \hat{x} > 0 \\ \lvert x - \alpha(e^{x+x\Delta} - 1)\rvert, & x \geq 0\ \&\ \hat{x} < 0 \\ \lvert\alpha(e^x - 1) - (x + x\Delta)\rvert, & x < 0\ \&\ \hat{x} \geq 0 \\ \lvert\alpha e^x(1 - e^{x\Delta})\rvert, & x < 0\ \&\ \hat{x} < 0 \end{cases}$$	Let $\alpha \in \mathbb{R}$ with $\alpha > 0$. With $\hat{x} = x(1+\delta)$ the relative error $\varepsilon = \frac{y_\Delta}{f(x)}$ is $$\varepsilon = \begin{cases} \lvert\delta\rvert, & x \geq 0\ \&\ \hat{x} > 0 \\ \frac{\lvert x - \alpha(e^{x(1+\delta)} - 1)\rvert}{x}, & x > 0\ \&\ \hat{x} < 0 \\ \frac{\lvert\alpha(e^x - 1) - x(1+\delta)\rvert}{\lvert\alpha(e^x - 1)\rvert} & x < 0\ \&\ \hat{x} \geq 0 \\ \frac{e^x\lvert 1 - e^{1+\delta}\rvert}{\lvert e^x - 1\rvert} & x < 0\ \&\ \hat{x} < 0 \end{cases}$$

TABLE I: Absolute and relative error bounds for common activation functions.

Softmax	Absolute Error Bound	Relative Error Bound
$s(x)_i$	For $\delta \in \mathbb{R}^+$, $\delta \leq \frac{1}{4}$ such that $\lvert\delta_i\rvert \leq \delta$ for all i in a vector $\underline{\delta}$, $$\lvert s(\underline{x} + \underline{\delta})_i - s(\underline{x})_i\rvert \leq 4.4445\bar{\delta}.$$	Let $\bar{\varepsilon} \in \mathbb{R}^+$, $\bar{\varepsilon} \leq \frac{1}{4}$. Let $\underline{x}, \underline{\varepsilon}$ and $\underline{\tilde{x}} \in \mathbb{R}^N$ such that $\forall\ i = 0, ..., N-1$, $\lvert\varepsilon_i\rvert \leq \bar{\varepsilon}$ and $\tilde{x}_i = x_i(1+\varepsilon_i)$. Let $\underline{y} = s(\underline{x})$ and $\underline{\tilde{y}} = s(\underline{\tilde{x}})$. $\forall\ i = 0, ..., N-1$, $$\left\lvert\frac{\tilde{y}_i - y_i}{y_i}\right\rvert \leq \lvert e^{x_{t'}\varepsilon_{t'} - x_t\varepsilon_t} - 1\rvert.$$ with t, t' such that $\forall\ (i, j) \in [0, N-1]^2$, $$\lvert e^{x_j\varepsilon_j - x_i\varepsilon_i} - 1\rvert \leq \lvert e^{x_{t'}\varepsilon_{t'} - x_t\varepsilon_t} - 1\rvert$$

TABLE II: Absolute and relative error bounds for the softmax function (vector-to-vector).

From Error Analysis to Uncertainty Propagation

- **Current direction**

Instead of propagating single values

$$x \rightarrow f(x)$$

we propagate interval representations of uncertainty

$$[x-,x+] \rightarrow f([x-,x+])$$

- Using:
  - interval partitioning,
  - polynomial approximations of nonlinear activation functions,
  - and bounded uncertainty propagation.

Towards Uncertainty-Aware Learning

- Traditional deep learning

input → Neural Network → Prediction

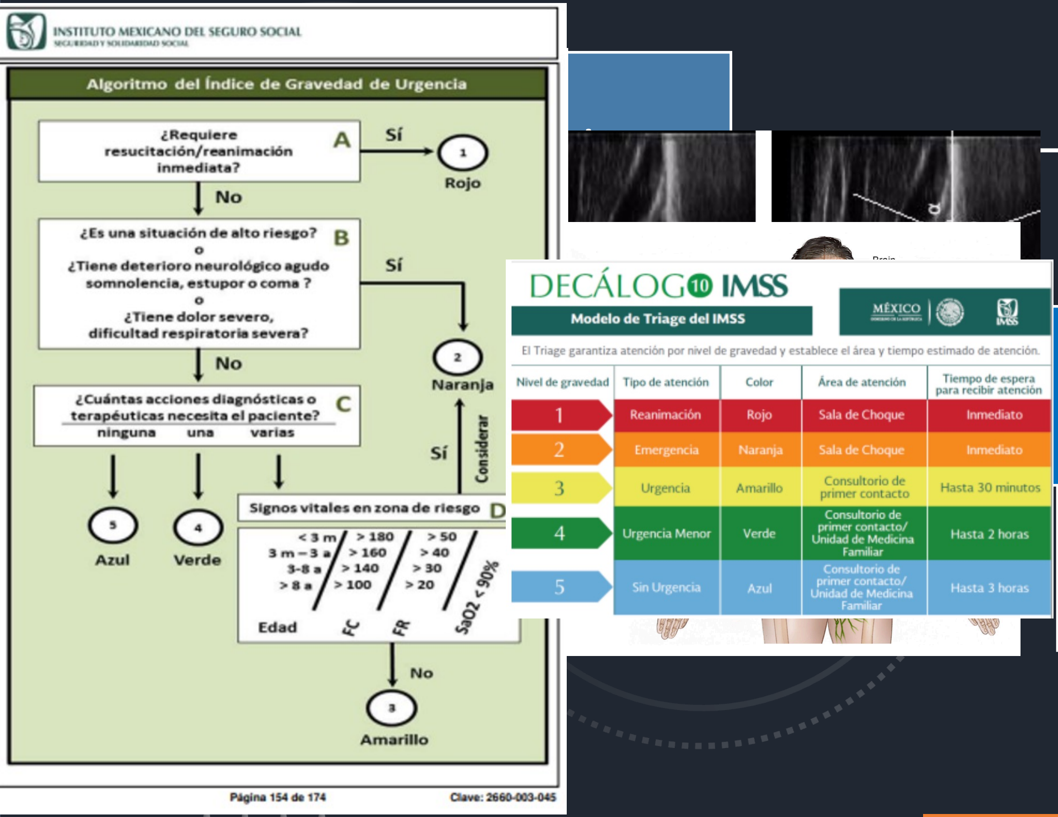
- Uncertainty-aware learning

input interval → Interval Network → Prediction+ uncertainty bound

Interval-Based Approach

- Guaranteed enclosures
- Safety over precision
- Integration with ML pipelines

One Computer Science Problem, Three Clinical Testbeds

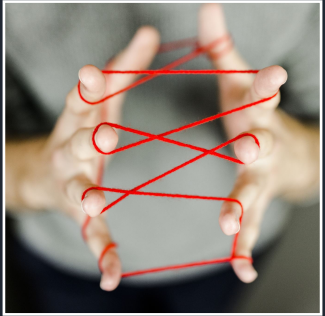


Clinical testbeds

- **DDH (Diagnosis)**
  - Imaging-based, operator variability, early screening
- **Relapsed NHL (Treatment)**
  - Structured data, multiple valid decisions, long-term outcomes
- **Hospital Triage (Risk Stratification)**
  - Resource allocation, partial information, time pressure

Why multiple domains?
Same uncertainty at different levels

Conclusion



The main contribution is not the application itself, but the framework for understanding numerical reliability in neural computation.

We should not only think about whether models are accurate, but also whether their computations are numerically stable and uncertainty-aware.

Merci
ありがとう
Thank You

I prefer true but imperfect knowledge, even if it leaves much indetermined and unpredictable, to a pretence of exact knowledge that is likely to be false.

-Friedrich Von Hayek,

Nobel Price lecture, 1974