

**The Attralucian Essays:**  
Exploring the Finite



First Edition

Copyright © 2026 by Kevin R. Haylett. All rights reserved.

This work is shared under the Creative Commons Licence.

Creative Commons CC BY-ND 4.0 License.

<https://creativecommons.org/licenses/by-nd/4.0/>

This work is intended for academic and research use. Any unauthorized distribution, modification, or commercial use beyond the creative use license is strictly prohibited. Typeset in

L<sup>A</sup>T<sub>E</sub>X

# The Attralucian Essays



The Learning and Generalization  
Problem: A Geofinitist Reinterpretation

Kevin R. Haylett

# The Learning and Generalization Problem: A Geofinitist Reinterpretation

## Overview

The problem of generalization lies at the heart of machine learning: how can a model trained on a finite dataset make reliable predictions on unseen data? Classical statistical learning theory addresses this through notions of risk, capacity, and convergence, often framed in asymptotic terms.

This paper does not reject these frameworks. Rather, it reinterprets generalization within a Geofinitist perspective, where all quantities are finite, measured, and accompanied by uncertainty. In this view, generalization is not a universal property of a model, but a *measured relation* between data provenance, model structure, resource constraints, and the region of prediction.

## From Classical to Measured Generalization

Let  $S = \{(x_i, y_i)\}_{i=1}^n$  be a finite dataset sampled from an unknown distribution  $\mathcal{D}$ , and let  $f$  be a model trained on  $S$ .

Classically, generalization is expressed via the gap between empirical and expected risk:

$$\mathcal{E}_{\text{gen}} = \mathbb{E}_{(x,y) \sim \mathcal{D}}[\ell(f(x), y)] - \frac{1}{n} \sum_{i=1}^n \ell(f(x_i), y_i).$$

This formulation depends on an unknown distribution and asymptotic reasoning. In practice, one observes only finite samples and finite computations.

Geofinitism reframes generalization as a *local, measurable property* within a finite data manifold.

## **Geofinitist Principles Applied**

### **1. Geometric Container Space**

Data does not occupy an abstract infinite space but lies on a structured manifold  $M$  with finite resolution. Generalization is meaningful only along trajectories within regions supported by data.

### **2. Measurement and Uncertainty**

Predictions are measured quantities:

$$\hat{y}(x) \pm \sigma_y(x),$$

where uncertainty arises from sampling, parameter estimation, and numerical computation.

### **3. Layered Representation**

In modern models, especially deep networks, representations evolve across layers. Generalization emerges from the aggregation of these transformations rather than a single global mapping.

### **4. Contextual Validity**

Generalization guarantees depend on the specific dataset, model, and training procedure. There is no universal guarantee independent of these factors.

### **5. Finite Constraints**

All learning occurs under finite resources: sample size, computational budget, and precision. These constraints define the limits of generalization.

## **Measured Local Generalization**

Let  $M$  denote an estimated data manifold with metric  $d_M$ . For a point  $x$ , define a local neighbourhood:

$$B_M(x, \rho) = \{x' : d_M(x, x') \leq \rho\}.$$

Define the local generalization measure:

$$G_M(x, \rho) = \mathbb{E}_{(x', y') \in B_M(x, \rho)}[\ell(f(x'), y')] - \mathbb{E}_{(x_i, y_i) \in S \cap B_M(x, \rho)}[\ell(f(x_i), y_i)] \pm \sigma$$

where  $\sigma_G$  captures uncertainty from finite sampling, model variability, and numerical effects.

## Uncertainty Model

A representative form:

$$\sigma_G(x, \rho) = k_1 \sqrt{\frac{1}{n_{\text{eff}}(x, \rho)}} + k_2 \|\nabla f(x)\| \rho + k_3 \epsilon_{\text{num}},$$

where:

- $n_{\text{eff}}$  is the effective sample count in the neighbourhood,
- $\|\nabla f(x)\|$  captures sensitivity,
- $\epsilon_{\text{num}}$  accounts for numerical precision.

## Out-of-Distribution Detection

Define the manifold distance:

$$d_{\text{OOD}}(x) = \min_{x_i \in \mathcal{S}} d_M(x, x_i).$$

Predictions are considered reliable only when:

$$d_{\text{OOD}}(x) \leq \rho_{\text{max}}.$$

Otherwise, the system reports abstention or elevated uncertainty.

## Layerwise Generalization

For a model with  $K$  layers:

$$G_{\mathbb{M}}(x) = \frac{1}{K} \sum_{\ell=1}^K G_{\ell}(x),$$

where each  $G_{\ell}$  measures local stability at layer  $\ell$ .

## Decision Rule

A prediction at  $x$  is accepted if:

$$|G_{\mathbb{M}}(x, \rho)| \leq \theta(x),$$

otherwise it is flagged as unreliable or out-of-distribution.

## Robustness via Perturbation

Define a perturbation operator  $P_{\eta}$  and smoothed generalization:

$$G_{\mathbb{M},\eta}(x) = \mathbb{E}[G_{\mathbb{M}}(P_{\eta}(x))].$$

Stable generalization requires bounded variation across  $\eta \in [\eta_{\min}, \eta_{\max}]$ .

## Formal Core (Geofinitist Generalization Protocol)

**Context.** Generalization is treated as a measurable, local property of a model relative to data and resources, not as a universal asymptotic guarantee.

**Measured System.** Define dataset provenance  $P_S$ , model provenance  $P_f$ , and training provenance  $P_{\text{Train}}$ . Predictions are measured outputs:

$$f(x) = (y, \varepsilon_y).$$

**Local Generalization.** For neighbourhood  $B_M(x, \rho)$ :

$$G_M(x, \rho) = E_{\text{test}} - E_{\text{train}} \pm \sigma_G.$$

**Reliability Criterion.** Accept prediction if:

$$|G_M(x, \rho)| \leq \theta.$$

**OOD Criterion.** Reject or abstain if:

$$d_M(x, S) > \rho_{\text{max}}.$$

**Robustness.** Evaluate stability under perturbations:

$$G_{M,\eta}(x).$$

**Reporting.** All results must include uncertainty bands and provenance.

**Collapse Note.** As  $n \xrightarrow{9} \infty$ ,  $\rho \rightarrow 0$ , and uncertainties vanish, this reduces to classical generalization theory (e.g. PAC bounds). Geofinitism retains

## **Interpretation**

Within this framework, generalization is not a binary property but a bounded, context-dependent quantity. It depends on:

- data density,
- model smoothness,
- computational precision,
- distance from observed regions.

## **Conclusion**

Geofinitism reframes generalization as a measurable property of finite systems. Rather than relying on asymptotic guarantees, it provides an operational protocol based on local geometry, uncertainty, and reproducibility.

This aligns learning theory with practice: models do not generalize universally, but within bounded regions where their predictions remain stable under finite perturbations.

Generalization, in this sense, is not the discovery of universal rules, but the maintenance of stability along measured trajectories within data-supported regions.