Taxonomic Networks: A Representation for Neuro-Symbolic Pairing

Georgia Tech College of Computing School of Interactive Computing

Zekun Wang[1], Ethan Haarer[1], Nicki Barari[2], Christopher MacLellan[1]

[1]Georgia Institute of Technology, [2]Drexel University

Drexel University

NeuS 2025

Key Takeaways

- **Unified Representation:** We introduce a single taxonomic hierarchy that serves both symbolic prototypes and neural soft trees, enabling seamless translation between paradigms.
- **Data Efficiency vs. Scalability:** Symbolic instantiation learns incrementally for maximal sample efficiency, while neural instantiation leverages batched training for high final accuracy.
- **Bidirectional Translation:** Closed-form mappings enables training symbolically then deploy neurally and vice versa.
- **Interpretability & Performance:** The shared taxonomy yields monosemantic concepts that are both human-interpretable and scalable with our neural approach.

Motivations

- **Dual Strengths of AI**: Symbolic methods offer interpretability and data efficiency; neural methods provide scalability and high asymptotic performance.
- **Gap in Representations**: Current neuro-symbolic framework use one of the approach as a sub-process of the other. However, no framework has reconciled them such that a single model and be instantiated with either neural or symbolic approach that supports transforms.
- **Our Solution:** A mode with shared representations that can be trained using either symbolic or neural approach via translation operators.

Methods

Taxonomic Networks are supervised hierarchical classification models. It realizes neuro-symbolic pair by building a single taxonomic representation that supports two learning paradigm: symbolic learning and neural learning

- **Shared representation**: A tree of hierarchical categories
- **Dual Updates:**
  - **Symbolic learning:** *Cobweb algorithm*. Incremental clustering via insert/merge/split/new-child operations determined by Categorical Utility measure on each node.
  - **Neural learning:** *Soft neural decision trees*. End-to-end soft decision routing of data trained by backpropagation.

Taxonomic Networks: an example of Neuro-Symbolic Paring

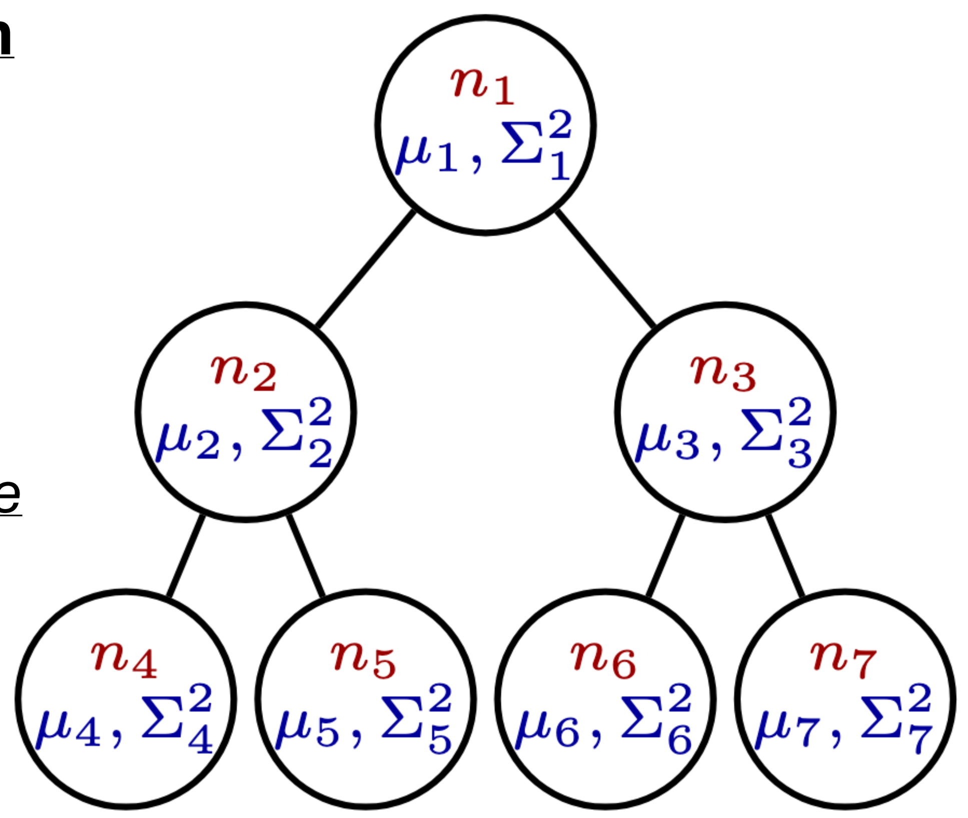
Symbolic instantiation

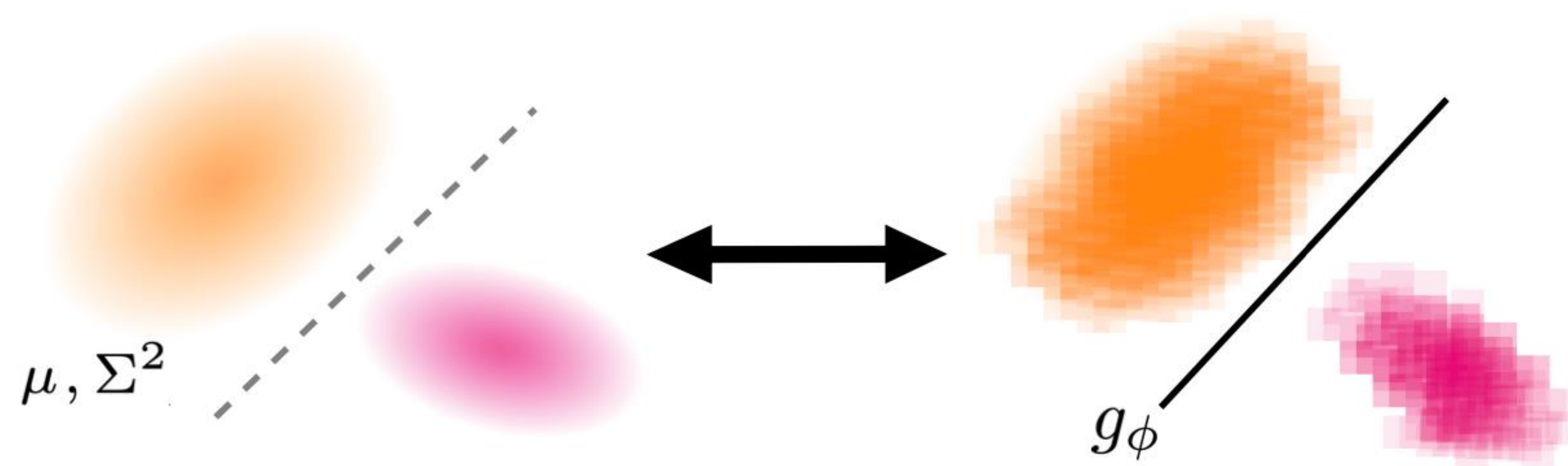
Each node stores:

- Data counts
- Gaussian params

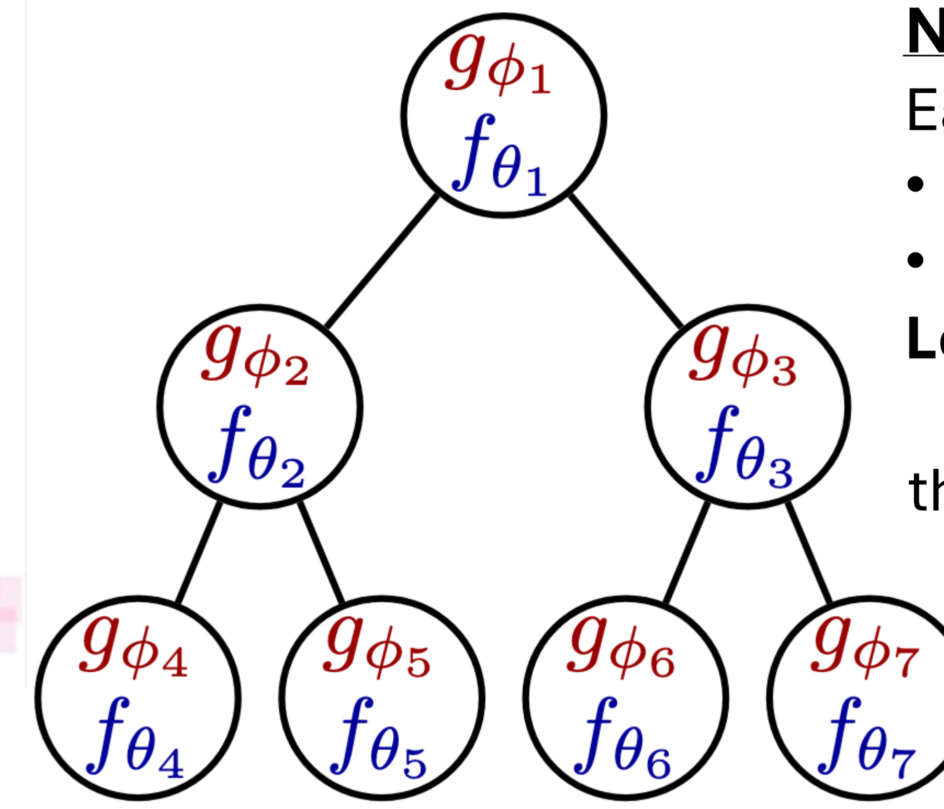
Learning objective:

Categorize data down the paths using *one of the four tree operations* (`insert, merge, add, create`) that maximize *Categorical Utility*





Transition Operators



Neural instantiation

Each node stores:

- Gating network
- Classification network

Learning objective:

Cross-entropy loss over the classification network outputs *weighted by the path probability* given By the gating network

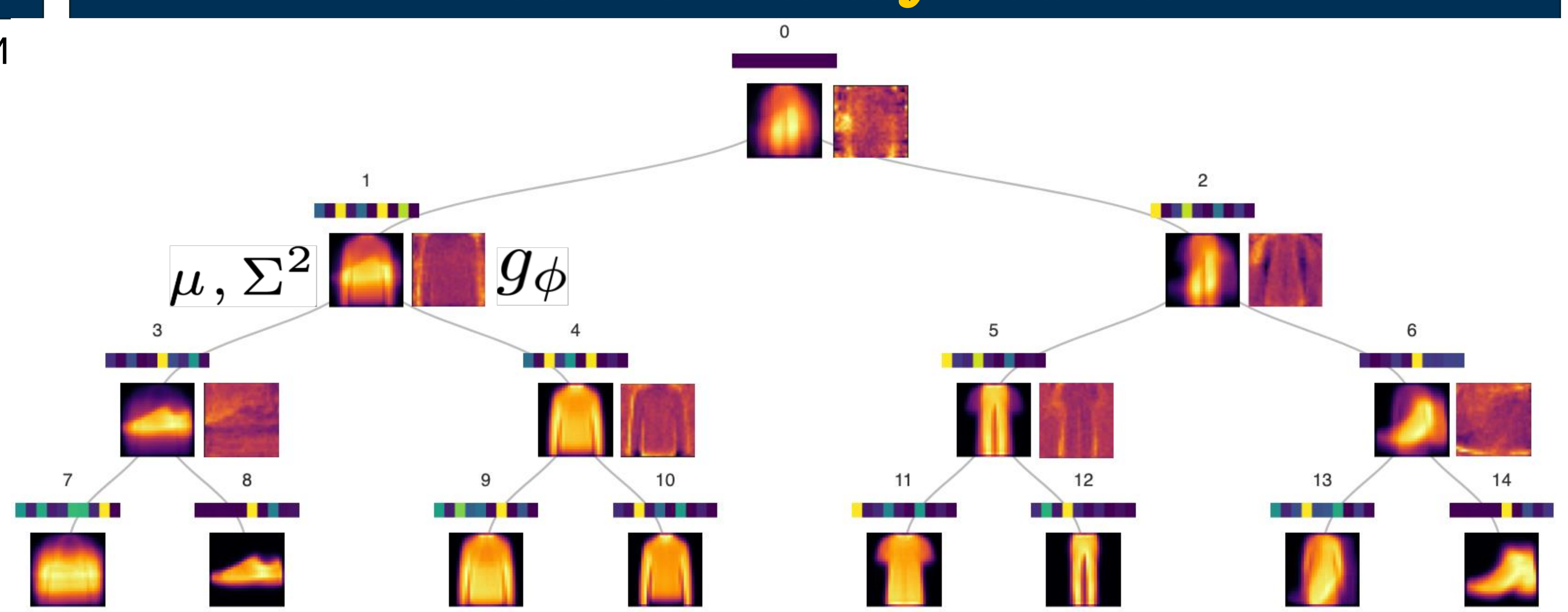
Experiments

Dataset: NMIST, FashionMNIST, and CIFAR-10

Metric: classification accuracy

- Analysis the learning curves for image classification on all datasets
- Inspect both Gaussian centroids and gating functions over the shared hierarchy

Learned Taxonomy Structure



Learning Curves