Lecture #14

Carnegie Mellon University

Special Topics:

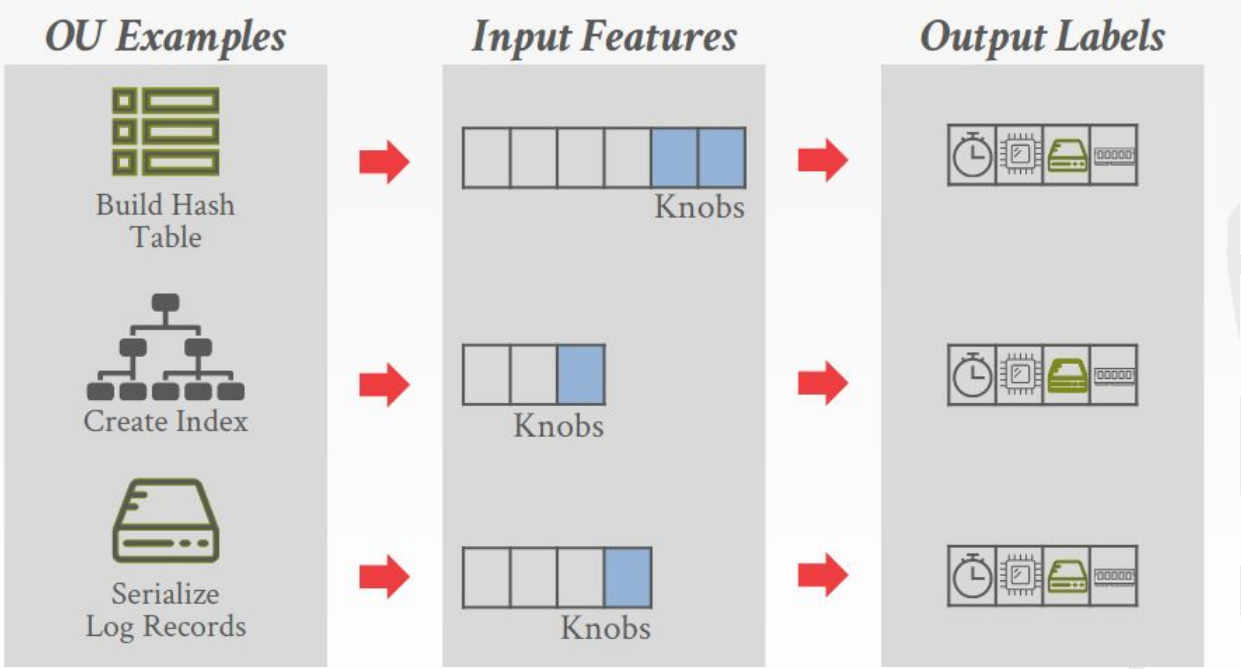
Self-Driving Database Management Systems

Behavior Modeling II

@William Zhang // 15-799 // Spring 2022

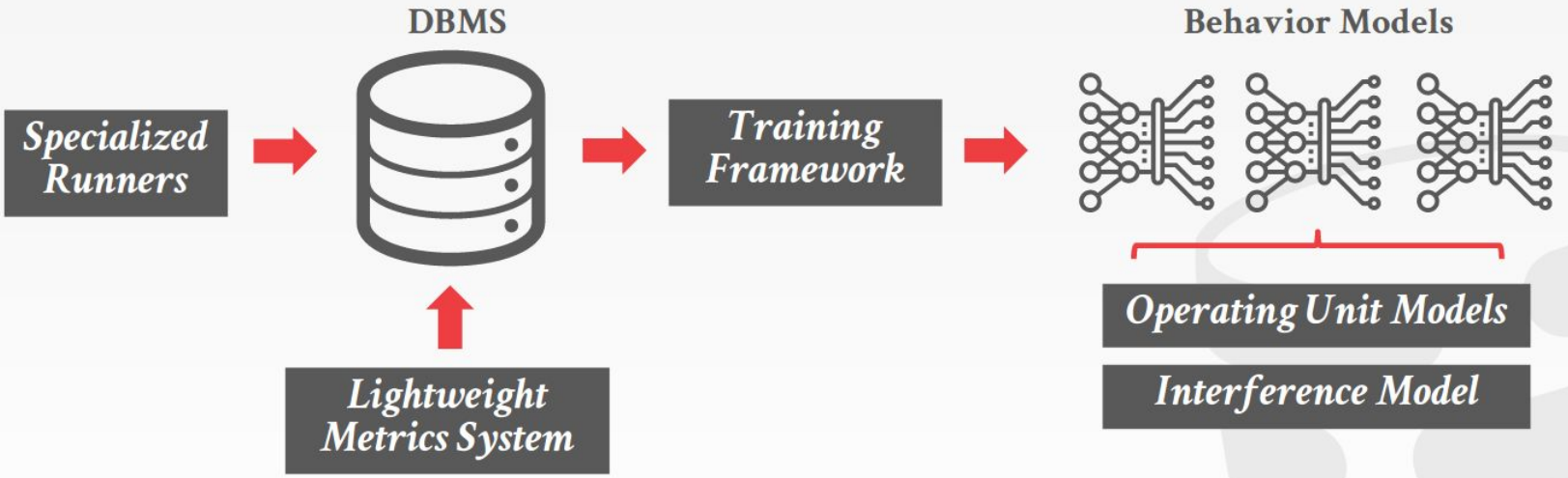
LAST CLASS (MB2)

Model via independent operating units



MB2: DECOMPOSED BEHAVIOR MODELING FOR SELF-DRIVING DATABASE MANAGEMENT SYSTEMS
SIGMOD 2021

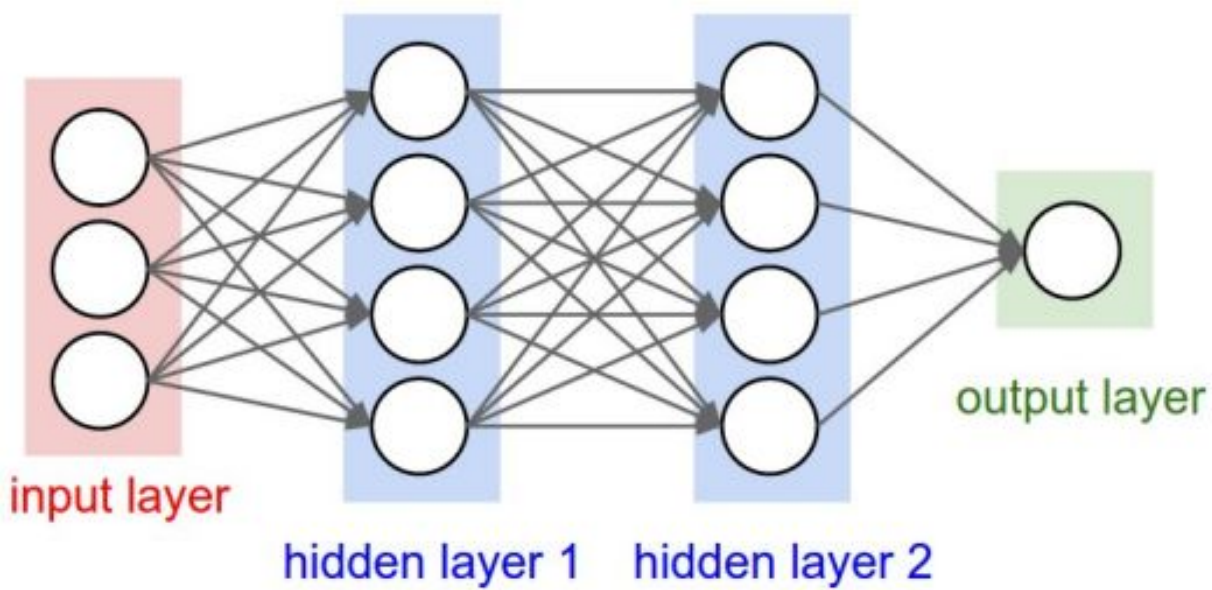
LAST CLASS (MB2)



MB2: DECOMPOSED BEHAVIOR MODELING FOR SELF-DRIVING DATABASE MANAGEMENT SYSTEMS
SIGMOD 2021

BACKGROUND

DNN Models



CHALLENGES

Query Plan Structure

Diverse Operator Features

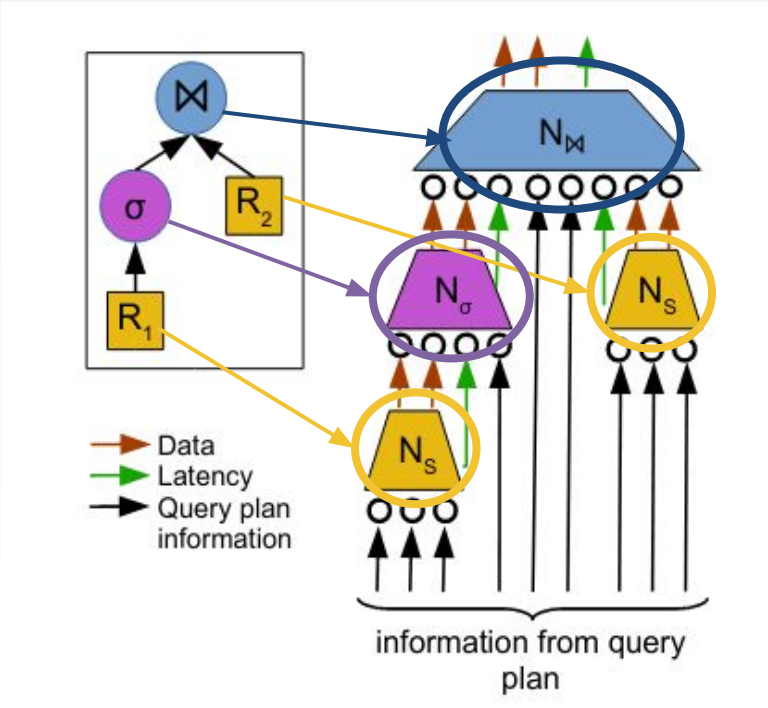
Position Independent Operator Behavior

QPPNet

Plan Structured Neural Network

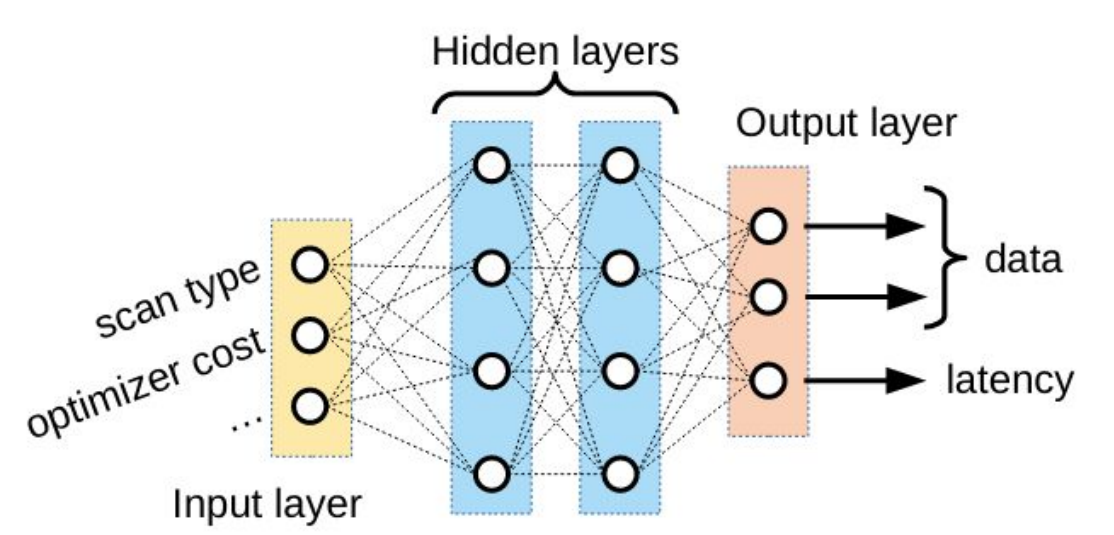
→ Operator Neural Units

→ Tree of Operator Neural Units



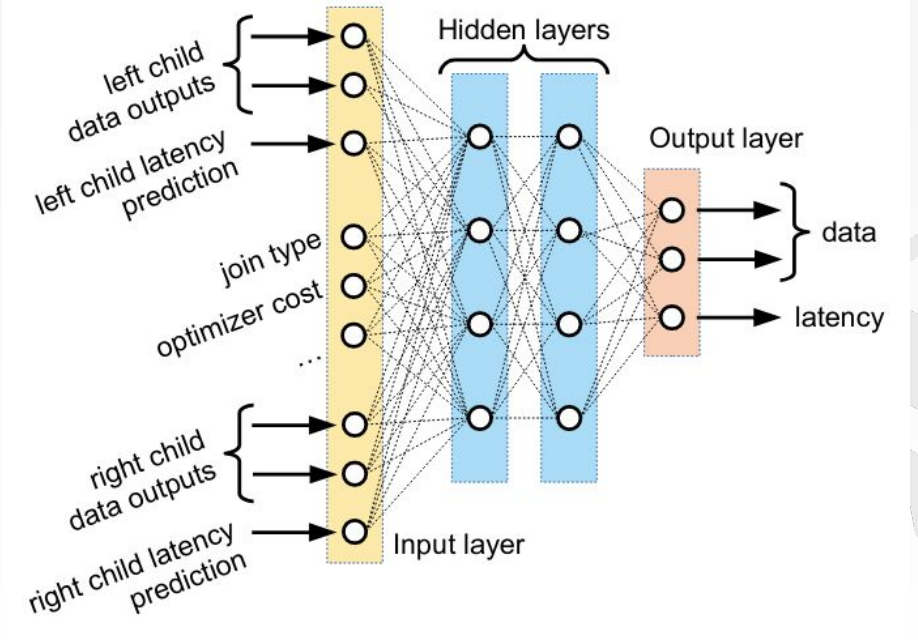
OPERATOR NEURAL UNIT

- Captures a distinct operator
- Outputs latency prediction and opaque vector



OPERATOR NEURAL UNIT (NON-LEAF)

- Incorporates outputs of child operators



OPERATOR INPUTS

Feature	PostgreSQL operators	Encoding	Description
Plan Width	All	Numeric	Optimizer's estimate of the width of each output row
Plan Rows	All	Numeric	Optimizer's estimate of the cardinality of the output of the operator
Plan Buffers	All	Numeric	Optimizer's estimate of the memory requirements of an operator
Estimated I/Os	All	Numeric	Optimizer's estimate of the number of I/Os performed
Total Cost	All	Numeric	Optimizer's cost estimate for this operator, plus the subtree
Join Type	Joins	One-hot	One of: semi, inner, anti, full
Parent Relationship	Joins	One-hot	When the child of a join. One of: inner, outer, subquery
Hash Buckets	Hash	Numeric	# hash buckets for hashing
Hash Algorithm	Hash	One-hot	Hashing algorithm used
Sort Key	Sort	One-hot	Key for sort operator
Sort Method	Sort	One-hat	Sorting algorithm, e.g. "quicksort", "top-N heapsort", "external sort"
Relation Name	All Scans	One-hot	Base relation of the leaf
Attribute Mins	All Scans	Numeric	Vector of minimum values for relevant attributes
Attribute Medians	All Scans	Numeric	Vector of median values for relevant attributes
Attribute Maxs	All Scans	Numeric	Vector of maximum values for relevant attributes
Index Name	Index Scans	One-hot	Name of index
Scan Direction	Index Scans	Boolean	Direction to read the index (forward or backwards)
Strategy	Aggregates	One-hot	One of: plain, sorted, hashed
Partial Mode	Aggregate	Boolean	Eligible to participate in parallel aggregation
Operator	Aggregate	One-hot	The aggregation to perform, e.g. `max, min, avg`

Table 2: QPP Net Inputs

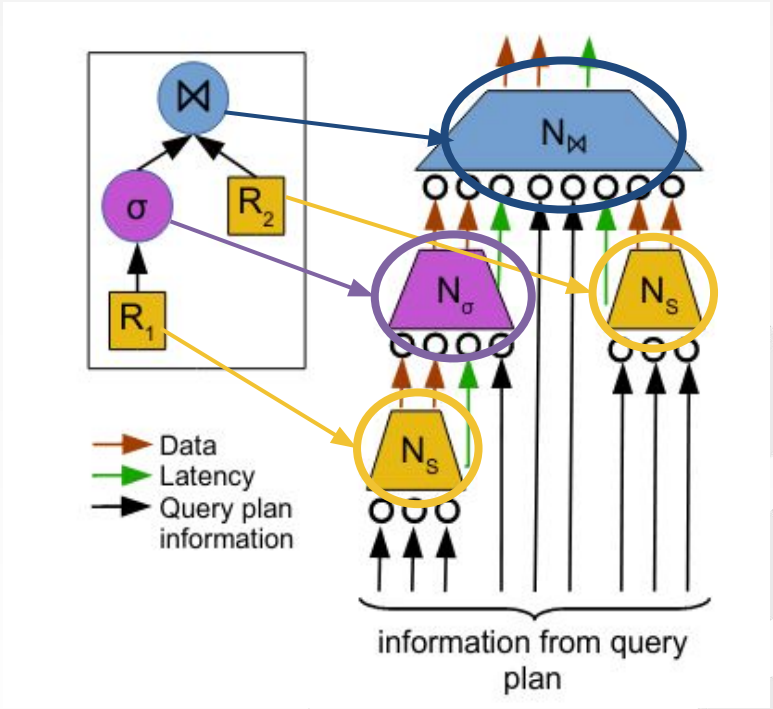
TREE OF NEURAL UNITS

Insights

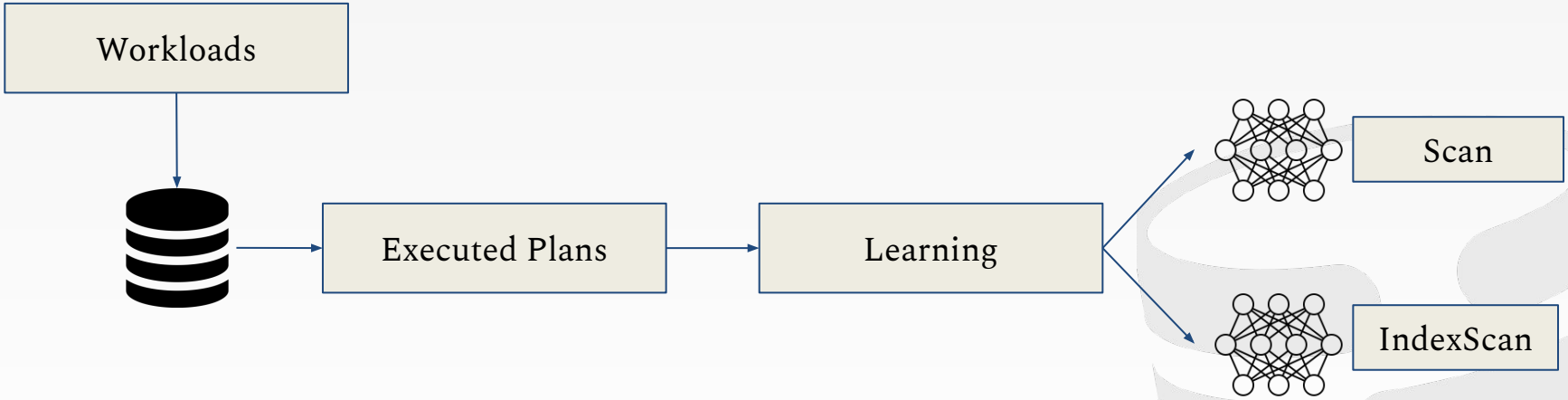
→ Non-independence (child → all ancestors)

→ Handle diverse operator input features

→ Single "Unit" per operator



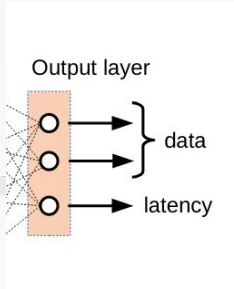
TRAINING PIPELINE

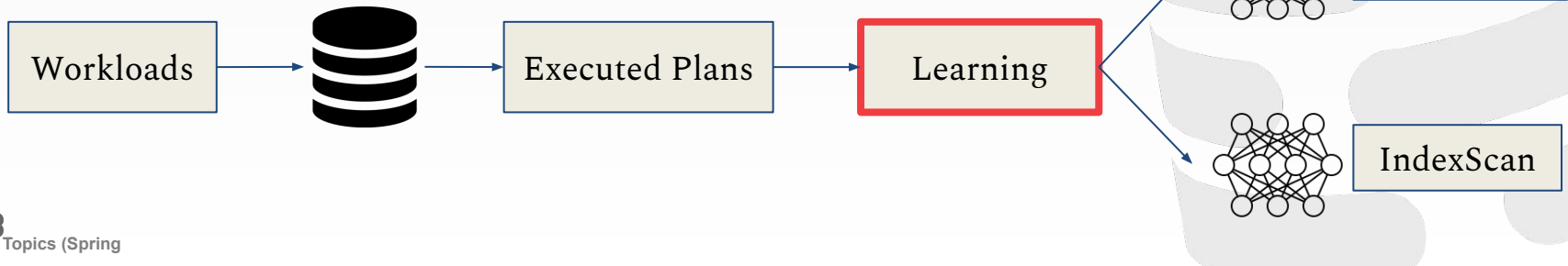


"LEARNING"

Minimize Latency Error

- Across all operators simultaneously
- Neural Unit Data Vector is "free-floating"
- Enables modeling dependence





EVALUATION

Workloads

→ TPC-H, TPC-DS (SF of 100GB)
→ 22 TPC-H templates, 69 TPC-DS templates
→ 22K queries

Train/Test Split

→ TPC-H 90%/10% data split
→ TPC-DS withhold 10 query templates
→ With all optimizations, takes ~24 hours

EVALUATION (PREDICTION)

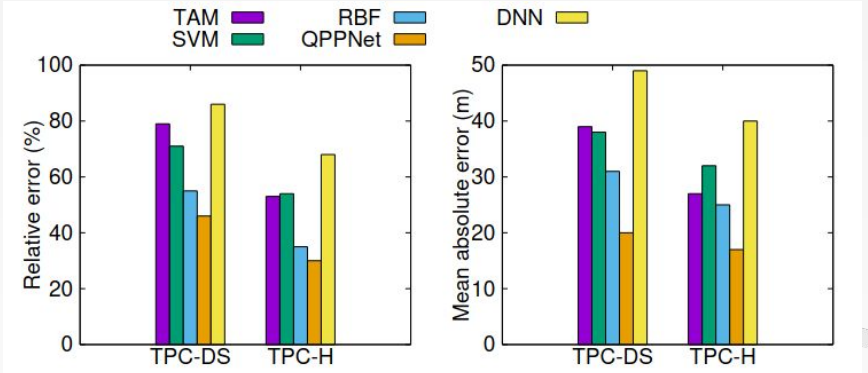
Baselines

→ SVM w. hand-picked features

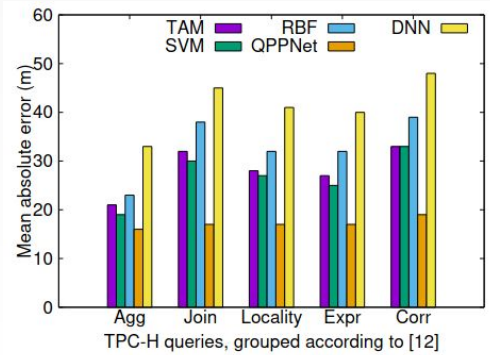
→ RBF w. hand-picked features

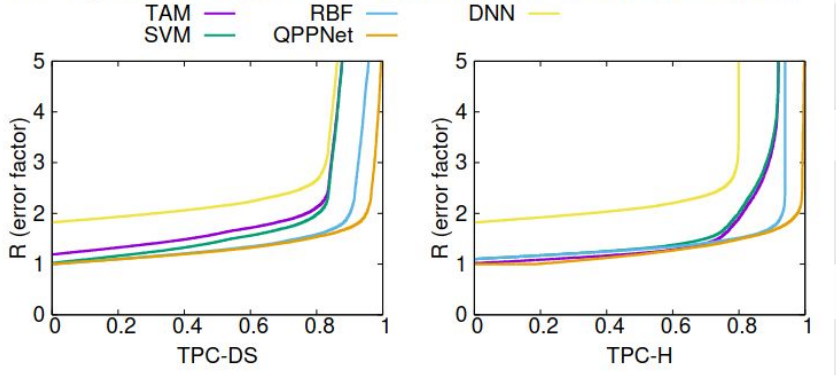
→ TAM ("tuned cost model")

→ DNN (non-tree deep neural network)

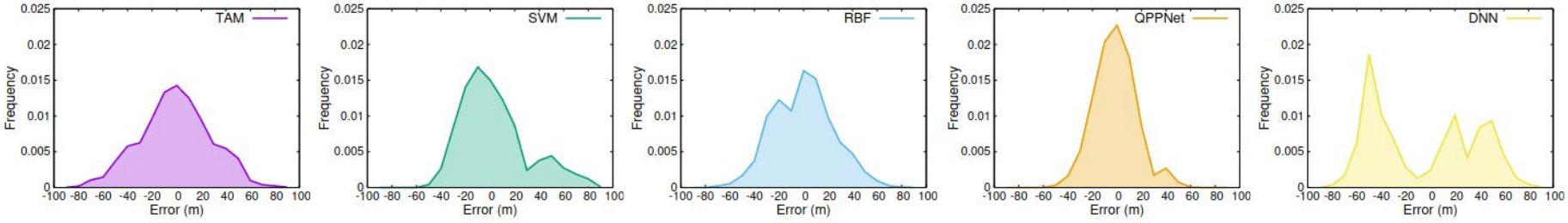


(a) Comparison of relative and mean absolute error for different workloads





EVALUATION (ERROR DISTRIBUTION)



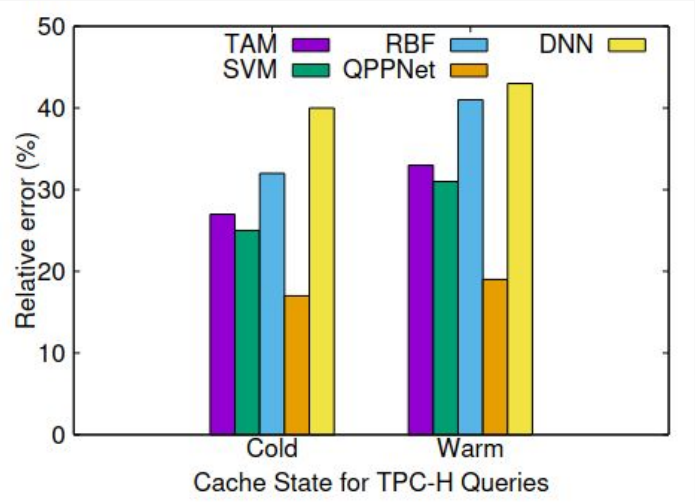
(a) TAM (b) SVM (c) RBF (d) QPPNet (e) DNN

Figure 7: Prediction error distributions for TPC-DS

→ Normalized to sum to 1
→ QPPNet narrow, center error bounds

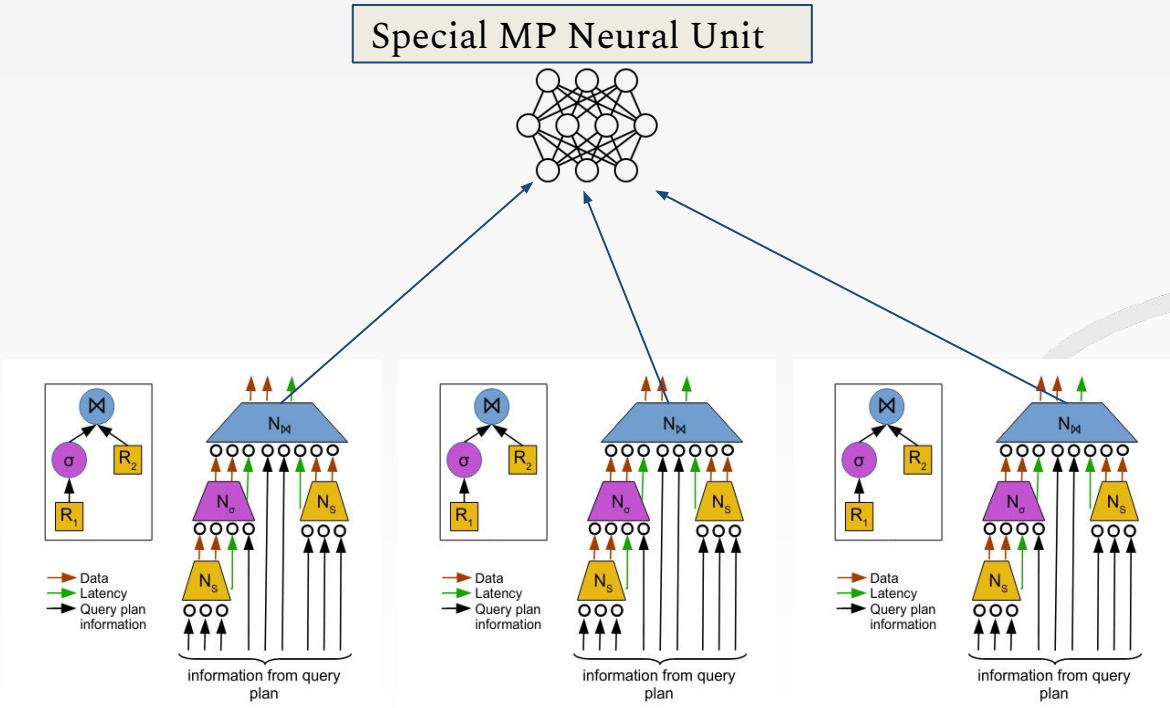
EVALUATION (CACHE)

→ Execution data from cold/warm cache
→ Cache state is not an input to model
→ QPPNet Error: Cold (17%), Warm (19%)



(b) Prediction error for warm vs cold cache

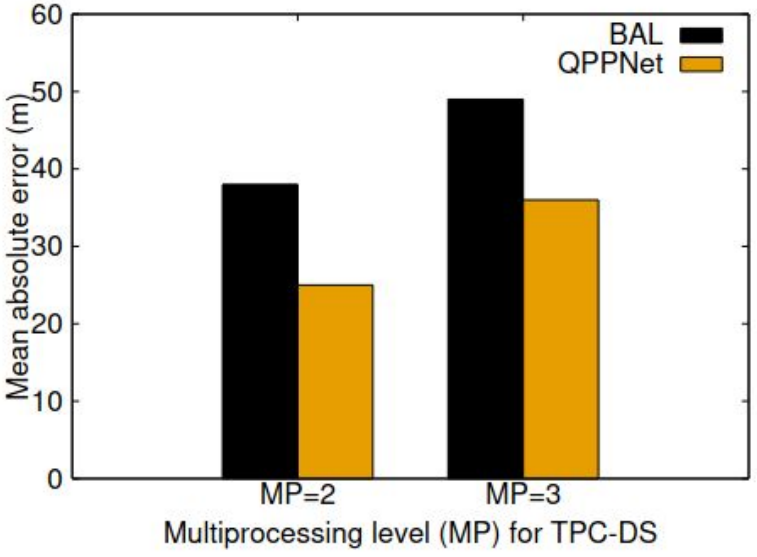
EVALUATION (PARALLEL)



Concurrent Query #1 Concurrent Query #2 Concurrent Query #3

EVALUATION (PARALLEL)

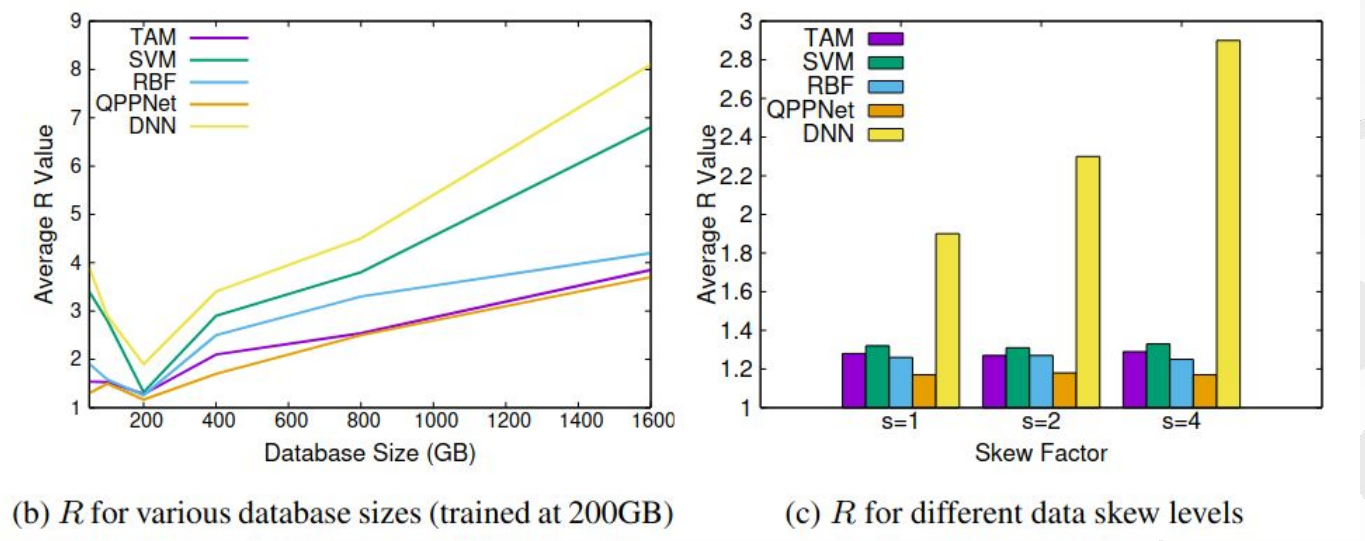
→ TPC-DS (2346 pairs, 52394 triples)
→ Comparison: BAL (SOTA predict buffer access latency)



(c) Prediction error per concurrency levels

EVALUATION (DATABASE SIZE)

→ Train at 1 DB size and test at other sizes (shrink v. expand)
→ Resistant to skew <u>but skew transfer not tested</u>



(b) R for various database sizes (trained at 200GB)

(c) R for different data skew levels

TAKEAWAYS

Plan Structured Neural Networks

→ Capture plan structure and operator interactions
→ Formulate complex (opaque) relationships between inputs

Comments

→ Workload "embedding" into the models
→ Questionable generalizability (or at least DB size)
→ "Interference Model"-approach to parallel
→ No direct incorporation of cache state
→ Targeting OLAP (TPC-H, TPC-DS)