Fully Transparent Machine Learning: Exact Factor Table Representation of GBMs

Lucas Muzynoski
Avenue Analytics
Insurance Data Science Conference 2025

Research Problem

- In the insurance industry, we demand accurate predictive models, but they often need to be transparent and interpretable to be useful in production.

Transparency: The underlying formula for the model is fully shared.
Interpretability: The model's calculations are reasonably understandable - stakeholders can see how it works and what factors drive predictions.

- Stakeholders and regulators need to know exactly how a model works before endorsing its use.

Existing Tradeoffs

- GLMs: Transparent (standard for regulatory filings) but limited predictive power without extensive feature engineering and regularization.
- GBMs: High accuracy but historically a 'black box', posing challenges for regulatory approval and internal buy-in.
- Post-hoc methods: Provide local explanations without full model transparency.
- Distillation: Approximations with some information loss.
- EBMs: Transparent, but limited interaction search and does not promote sparsity.

A Black Box into a Glass Box:

The Challenge: How can we unlock the power of gradient boosting while maintaining complete transparency?

Making it work:

- Start with a high-performance GBM like LightGBM
- Create methods for converting and consolidating trees into factor tables, exactly
- Use novel regularization parameters for sparsity and interpretability
- Multi-objective tuning to balance performance and interpretability

Result:

- Factor tables! A GBM that is now transparent and interpretable
- Predictions match exactly with underlying LightGBM model
- Performance parity with state-of-the-art methods

How It Works: Tree to Factor Table

Simple Example:

```
Tree:
Age
|-- <=30: 0.2
|-- >30
    |-- Sedan: 0.5
    |-- Other
        |-- SUV: -0.1
        |-- Truck: 0.3
```

Factor Table:

Age	Vehicle	Factor
<=30	*	0.2
>30	Sedan	0.5
>30	SUV	-0.1
>30	Truck	0.3

Key Point: Every tree can be converted to factor tables with *zero* information loss

Note: Every tree can be decomposed into individual node contributions, enabling us to separate main effects from interactions during consolidation

Consolidating for Interpretability

A GBM with 100+ trees would create 100+ factor tables - too many to be interpretable. We consolidate by decomposing each tree node into mini factor tables, then combining them strategically.

Two Consolidation Strategies:

- **ANOVA-style:** Separate main effects from interactions (better interpretability)
- **Full consolidation:** Combine any tables where features are subsets (more compact)

Both preserve exact prediction equivalence with the underlying GBM!

Outcome: A compact set of interpretable factor tables (often 1-10) instead of 100+ trees

From GBM Ensemble to Consolidated Tables

1. GBM Ensemble (e.g., 100s of decision trees)

->

2. Individual Tree Tables Each tree is an exact factor table.

->

3. Consolidated Factor Tables Equivalent to original GBM!

Complexity Control Methods

Research Challenge: Even with consolidation, a standard GBM can produce excessive numbers of factor tables, ruining interpretability

Complexity Control Methods

Research Challenge: Even with consolidation, a standard GBM can still produce too many factor tables

Our Solution: New regularization penalties that discourage the model from using new feature combinations

How it works:

- Penalize splits that introduce feature combinations not seen before in the ensemble
- Penalize splits that add new feature combinations within individual trees
- Automatically promotes sparsity and interpretability

Outcome: Models automatically select only a small number of features and interactions

Tuning GBMs for Performance and Interpretability

Optimization Problem: $\min_{\theta \in \Theta}\{-CV(\theta), C(\theta)\}$

Two Objective Functions:

- **Performance:** Cross-validation accuracy
- **Complexity:** Median number of consolidated factor tables

Outcome: A Pareto frontier of optimal hyperparameters. Choose between maximimum performance or more interpretability.

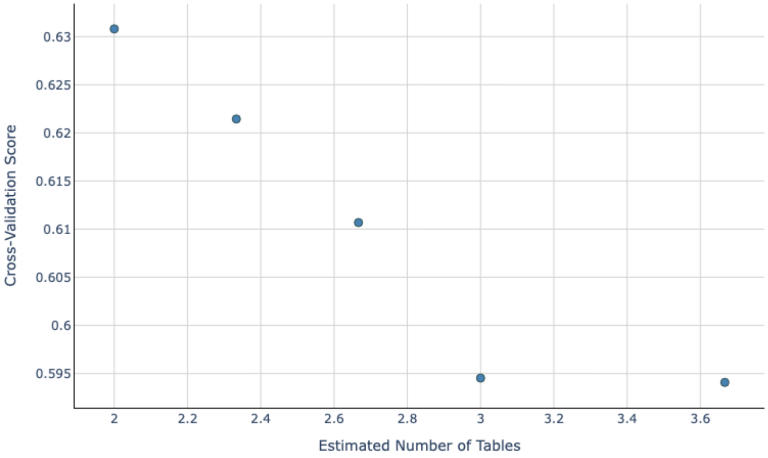

Figure 1: Pareto Frontier

Case Study 1: Recidivism Prediction

Dataset: Broward County recidivism data (ProPublica)
Task: Binary classification (2-year recidivism)
Benchmark: Proprietary COMPAS algorithm
Key Result: 123-tree LightGBM ensemble converted to 2 factor tables (zero information loss)
Experimental Results:

Method	AUC
COMPAS	0.696
Random Forest	0.676
EBM	0.728
Our Method	**0.726**

Finding: Competitive performance with complete model transparency

Recidivism: The Complete Model

Intercept: -0.185

Table 1: Prior Charges × Sex (showing first 6 rows)

Prior Charges	Sex	Factor
0	Female	-1.100
$<=$1.5	Female	-0.496
$<=$2.5	Female	0.044
0	Male	-0.775
$<=$1.5	Male	-0.321
$<=$2.5	Male	0.097

22 total rows

Recidivism: Age Effects

Table 2: Age × Sex (showing first 6 rows)

Sex	Age	Factor
Female	$<=20.5$	1.417
Female	$<=21.5$	1.082
Female	$<=22.5$	0.857
Male	$<=20.5$	1.470
Male	$<=21.5$	1.136
Male	$<=22.5$	0.911

26 total rows

Manual Prediction: Look up 2 numbers, add them to the intercept, then apply the logistic function

Case Study 2: Insurance Claim Frequency

Dataset: French Motor MTPL (678,013 policies)

Task: Poisson regression for claim frequency

Performance Comparison:

Method	Poisson Deviance
Random Forest (100 trees)	0.690
EBM	0.599
Our Method (Interpretable)	**0.593**
Our Method (Best Perf.)	**0.583**

Insurance Factor Tables: Structure

Model Configuration (Interpretable Variant):

- **Baseline:** -2.302

Component	Feature Set	Structure	Parameters
Table 1	Vehicle Age $\times$ Gas Type	VehAge, VehGas, Factor	50
Table 2	Driver Age	DrivAge, Factor	25
Table 3	Bonus-Malus $\times$ Vehicle Age	BonusMalus, VehAge, Factor	400+

Key Advantages & Applications

What This Enables:

- **Regulatory Compliance:** Factor tables ready for filing, familiar to regulators
- **Full Transparency:** Understand exactly how predictions are made while retaining GBM performance
- **Plug-and-Play:** Integrates directly into existing rating engines and actuarial software

Perfect For:

- **Insurance Pricing & High-Stakes Finance:** Where regulatory approval is required
- **Healthcare & Criminal Justice:** Where algorithmic transparency is mandatory
- **High-Dimensional Problems:** Built-in feature selection for true sparsity

Bottom Line: Deploy state-of-the-art ML where transparency and interpretability is required

Questions?

Contact: Email: info@avenue-analytics.com LinkedIn: Lucas Muzynoski

Link to pre-print of full paper