Kumo

USE CASE · LEAD SCORING

Lead Scoring Powered by a Relational Foundation Model

Which leads will convert is a million-dollar question - and most companies answer it badly. Hiring ML teams means months of feature engineering from flat tables. LLMs tokenize your CRM data without understanding relationships. KumoRFM is the world's first foundation model for relational data.

3.2x
MORE ACCURATE
Than traditional ML on lead conversion

<1 hr
TO PRODUCTION
From raw tables to deployed scores

85%+
AUROC
On real-world B2B lead conversion

0
MANUAL FEATURES
Discovered automatically from relational data

EXECUTIVE SUMMARY

The opportunity

Which leads will convert is a million-dollar question - and most companies answer it badly. Hiring ML teams means months of feature engineering from flat tables. LLMs tokenize your CRM data without understanding relationships. KumoRFM is the world's first foundation model for relational data.

3.2x

More accurate

Than traditional ML on lead conversion

<1 hr

To production

From raw tables to deployed scores

85%+

AUROC

On real-world B2B lead conversion

0

Manual features

Discovered automatically from relational data

TRUSTED BY INDUSTRY LEADERS

Databricks Snowflake DoorDash Reddit Sainsbury's Coinbase Discord iFood Ro

Chime Yieldmo Catalina Faire Tubi Quizlet

LEADERSHIP

Built by pioneers in AI

Leadership from Airbnb, Stanford, LinkedIn, and Pinterest. Creators of PyG, the #1 graph ML library, and 40+ peer-reviewed papers at NeurIPS, ICML, and KDD.

Vanja Josifovski

CEO & Co-Founder

Former CTO at Airbnb and Pinterest. 20+ years leading ML at scale.

Jure Leskovec

Co-Founder & Chief Scientist

Stanford Professor. Co-creator of PyG (#1 Graph ML library) and Relational Deep Learning.

Hema Raghavan

Co-Founder & Head of Engineering

Former Sr. Director of Engineering at LinkedIn. Built ML infrastructure serving 900M+ users.

BACKED BY

Sequoia Capital

Led by Konstantine Buhler & Bill Coughran

Carl Eschenbach

CEO of Workday · Board Advisor

40+
Peer-reviewed papers
NeurIPS, ICML, KDD

#1
Graph ML library
PyG (PyTorch Geometric)

SOC 2
Type II certified
Enterprise governance

THE CHALLENGE

Two approaches to lead scoring. Neither works.

01 Hire an ML team: They spend months hand-crafting features from flat tables - company size, industry, last email open. This captures a fraction of the signal. Cost: $50K-$1M+ per model, 6-12 weeks, and 53-88% never reach production (Gartner, IDC).

02 Try LLMs: They tokenize your CRM data without understanding the relationships between tables. An LLM sees text, not the graph of contacts, accounts, opportunities, and interactions that actually predicts conversion.

03 Both approaches miss what matters most: the relational structure of your data. The connections between customers, products, support tickets, and billing events contain the richest predictive signals.

04 Meanwhile, sales reps learn to disregard scores within weeks. The models that do ship degrade quickly - with ~30% annual maintenance cost.

> *53–88% of ML models never reach production. The ones that do degrade quickly — with ~30% annual maintenance cost.*
>
> Sources: Gartner, IDC, Dimension Research

80%
Of ML time on feature engineering
Source: Cognilytica

$1M+
Cost per custom ML use case
Source: McKinsey

3–5
Models shipped per year
Source: McKinsey

THE SOLUTION

KumoRFM: the world's first foundation model for relational data.

1. KumoRFM connects directly to your data warehouse and learns from the relationships across CRM, product usage, support, and billing tables. No feature engineering. No flat tables.
2. The Graph Transformer attends across multiple columns, tables, and hops - discovering patterns that no ML team or LLM would find.
3. 85%+ AUROC on real-world lead conversion tasks - beating both traditional ML and LLM-based approaches (RelBench validated). Superhuman accuracy out of the box.
4. Setup takes under an hour. Kumo Enterprise allows fine-tuning for even higher accuracy. Continuous retraining keeps scores accurate as your data evolves.

HOW KUMO WORKS

From relational tables to deployed predictions

01

Connect your tables

Point Kumo at your CRM, product usage, marketing, support, and billing tables. Kumo infers the relational schema automatically.

02

Define the prediction

Predictive Query Language reads like SQL but describes future outcomes. No ML expertise required.

03

Deploy with explanations

Kumo trains a relational deep learning model, evaluates accuracy, and deploys to production with human-readable explanations.

DEPLOYED AT

Databricks Snowflake DoorDash Reddit Sainsbury's Coinbase Discord iFood Ro Chime

IN PRODUCTION TODAY

Customer results

Leading companies use Kumo to drive measurable business outcomes from their existing relational data.

5.4x conversion lift

Databricks

Lead-scoring models built on Kumo's relational deep learning delivered 5.4x improvement in lead conversion prediction accuracy.

Record accuracy lift

Reddit

Kumo graph embeddings drove Reddit's biggest-ever improvement in ad engagement models - a record-setting lift across their entire ML organization.

$100M+ GMV impact

DoorDash

Restaurant recommendations driving hundreds of millions in GMV. Expanded to notification reranking and send-time optimization.

Expansion revenue

Snowflake

Personalized solutions recommendations driving expansion revenue growth through relational data intelligence.

£10M+ per month

Sainsbury's

Search reranking models built in weeks. 7% improvement in search-to-purchase results.

+10% conversion

iFood

Enhanced recommendation accuracy across 80M monthly predictions, driving measurable revenue growth.

ALSO IN PRODUCTION AT

Coinbase Discord Ro Chime Yieldmo Catalina Faire Tubi

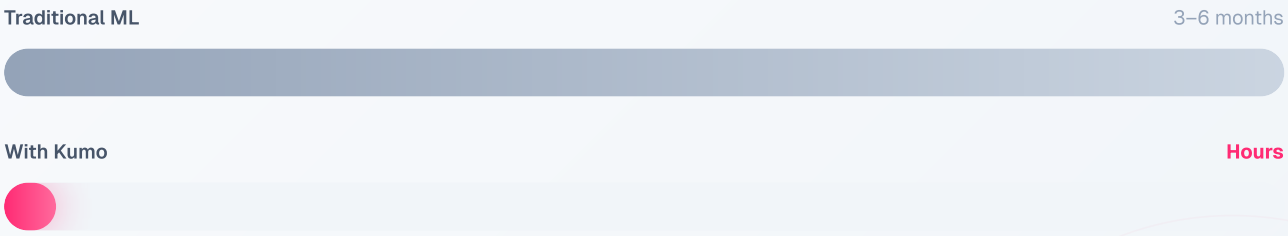
SIDE BY SIDE

Traditional ML vs. Kumo

Industry data from McKinsey, Gartner, and Dimension Research.

DIMENSION	TRADITIONAL ML	WITH KUMO
Input signals	50–100 hand-crafted features	**Entire relational database**
Setup time	6–12 weeks of engineering	**Under 1 hour**
Cost per model	$50K – $1M+ (McKinsey)	**Fraction - single platform**
AUROC	0.65–0.75	**0.85+ (RelBench validated)**
Explainability	Black box or basic SHAP	**Human-readable relational explanations**
Maintenance	~30% annual cost	**Continuous automatic retraining**

TIME TO PRODUCTION COMPARISON



UNDER THE HOOD

Why relational deep learning works

01

Relational deep learning discovers features automatically from your database schema - no manual feature engineering or flat-table extraction.

02

Graph Transformers attend across multiple tables and hops, learning patterns that span CRM, product, billing, and support data simultaneously.

03

Kumo's foundation model is pre-trained across hundreds of schemas, providing strong baselines even before fine-tuning on your data.

04

Every prediction includes human-readable explanations grounded in your actual data relationships - not abstract feature importances.

05

Continuous retraining ensures models stay accurate as your data evolves, eliminating the ~30% annual maintenance cost of traditional ML.

06

SOC 2 Type II certified with data residency controls, RBAC, audit logs, and approval workflows for enterprise governance.

STANFORD RELBENCH BENCHMARK

Superhuman accuracy on predictive tasks

Across 30 prediction tasks on 11 relational databases, Kumo outperforms both LLM-based approaches and expert PhD data scientists.

30	11	#1
Prediction tasks	Relational databases	Overall ranking

PLATFORM BREADTH

55+ use cases from a single platform

New customer acquisition

Lead scoring · Account scoring · Propensity to buy

Customer retention

Churn prediction · Engagement scoring · Win-back targeting

Product personalization

Recommendations · Search ranking · Content personalization

Next best action

Cross-sell / Upsell · Campaign optimization · Send-time optimization

Optimizing growth

Customer lifetime value · Revenue forecasting · Basket optimization

Fraud detection

Transaction fraud · Account takeover · Promo abuse

Entity resolution

Identity matching · Duplicate detection · Household mapping

Forecasting

Demand forecasting · Inventory planning · Seasonal trends

RESEARCH FOUNDATION

40+ peer-reviewed papers

ICML 2024 — **KumoRFM: A Relational Foundation Model for Predictive Analytics**
K. Huang, M. Fey, J. Leskovec et al.

NeurIPS 2024 — **Relational Deep Learning: Graph Representation Learning on Relational Databases**
M. Fey, W. Hu, K. Huang, J. Leskovec et al.

NeurIPS 2024 — **RelBench: A Benchmark for Deep Learning on Relational Databases**
J. Robinson, R. Miao, K. Huang et al.

Explore all 40+ publications at kumo.ai/research

Kumo

Your pipeline data already contains the conversion signal.

See what Kumo can predict from your existing relational database. Request a 30-minute demo - we'll show you results on your actual data.

Request a Demo | Try Kumo Free

WEBSITE	TRY KUMO FREE	CONTACT
kumo.ai	kumorfm.ai	kumo.ai/contact

Databricks Snowflake DoorDash Reddit Sainsbury's Coinbase Discord iFood Ro Chime Yieldmo Catalina