Creating and operating ML models from event-based data using a feature engine and a feature store

Davor Bonaci
CEO
KASKADA

Dr. Charna Parkey
VP, Product
KASKADA

Taimur Rashid
Chief Business Development Officer
redis

Photo by Wesley Tingey on Unsplash

Delivery Trends - Pandemic

- 30% drop in ordered items being found
- Average customer basket size +>35% month over month
- 500% increase in year over year order volume
- More localized events needed to predict which items will sell out
  - Increase in the number of items needing to be scored
  - Increase in the number of shops needing predictions
  - Noise reduction is critical

ML models can only be as good as the data we give it

Photo by [Jasmin Sessler](#) on [Unsplash](#)

Latency is an outage

500

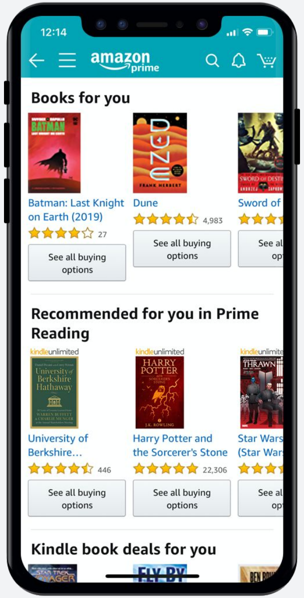
million

You need native time travel and a low latency, high throughput feature store

redis KASKADA

Feature Stores for ML

Time travel is hard



Personalization

Amazon: *Contributes 35% of revenue*

Recommendations

Netflix: *+13% in revenue due to savings*

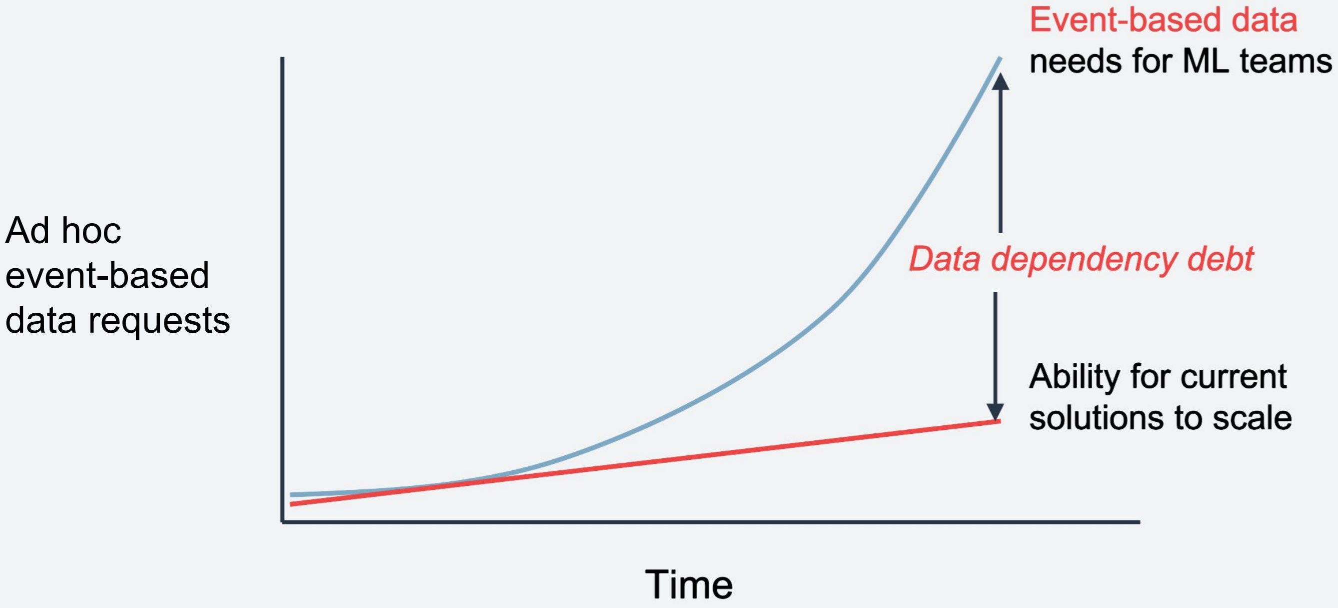
Fraud Alerts

FICO: *+30% in detection of CNP fraud*

redis KASKADA

Feature Stores for ML

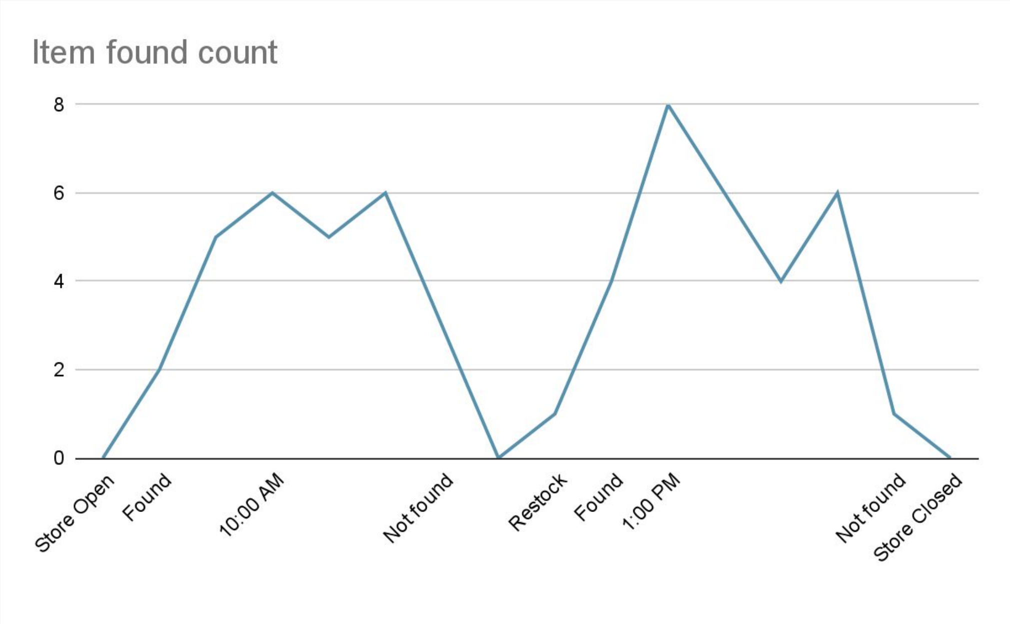
The growth in demand for unpredictable event-based data needs is increasing



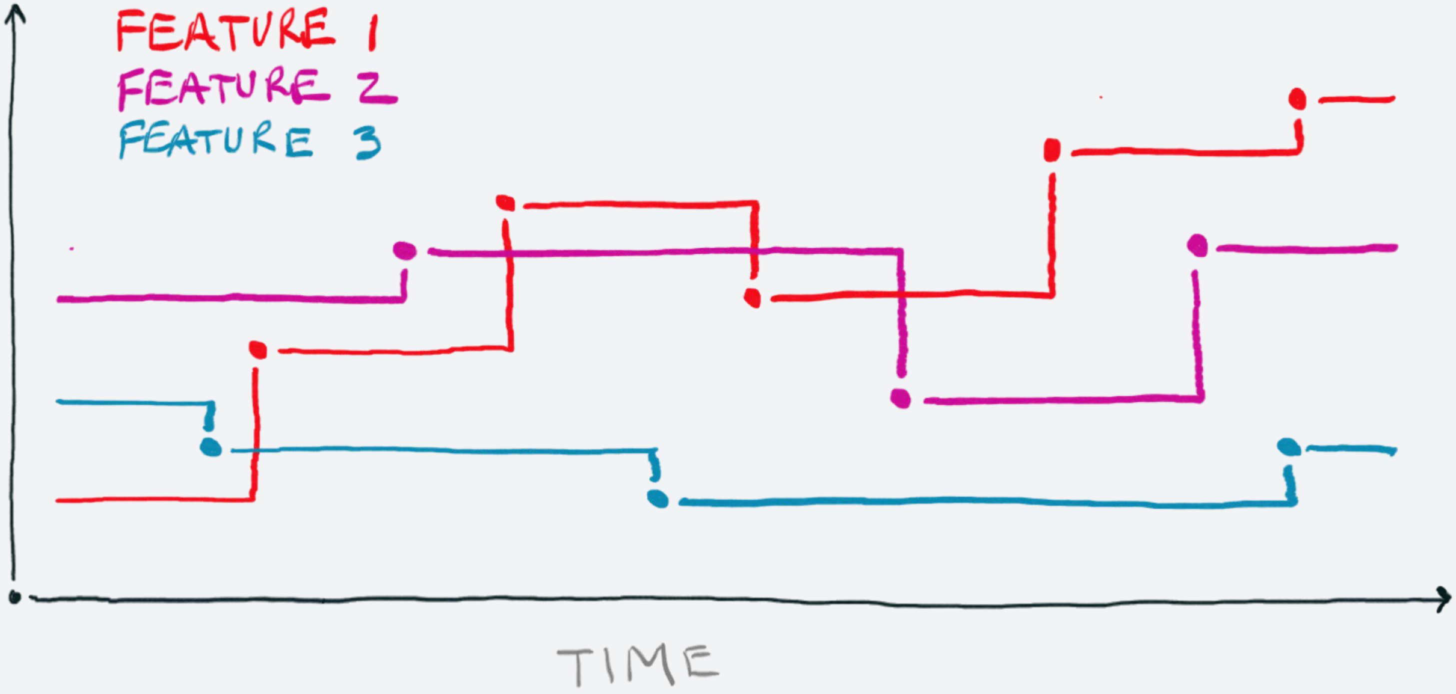
Instant Iteration Requires

- Historical feature value generation to try new features
- Expressive time selection to specify your model context iteratively
- Joining values between different entities, at precise times — without leakage
- Shared feature definitions to power live models

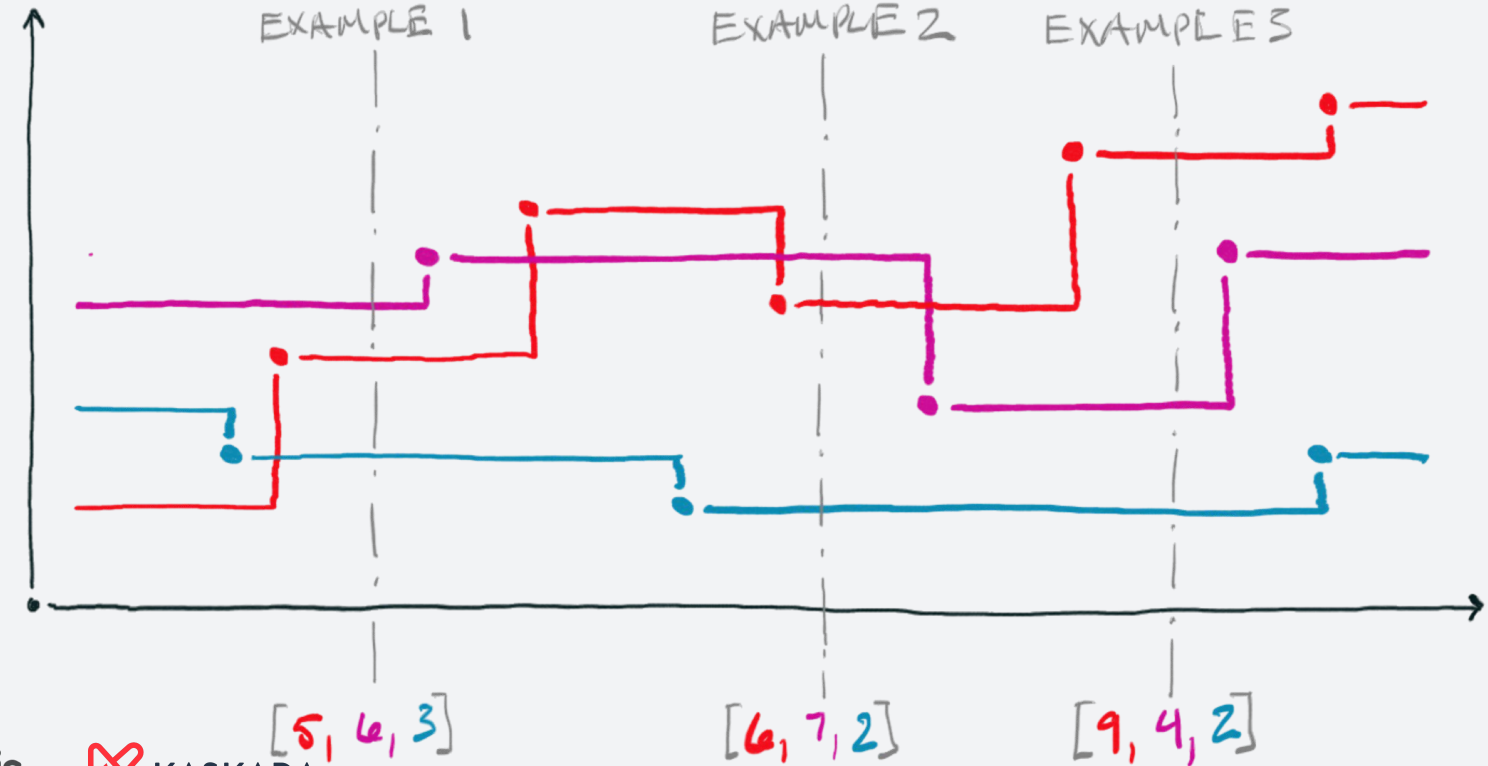
Relative event times are important



Feature definitions define *what* to compute



Time selection defines *when* to compute



redis

KASKADA

Feature Stores for ML

Discrete + continuous time temporal processing

Item Entity Features

- Historical purchase count
- Historical replacement rate
- Historical found rate
- Time since last found
- Expected time to next not found

Shopper Features

- Time of day shopping begins
- Day of week shopping begins

Retailer Entity Features

- Historical retailer availability
- Store location
- Restock times
- Store hours

Region Features

- Found rate of parent product category in the region

Photo by Oleksii S on Unsplash

Getting to production is hard for real-time inferencing

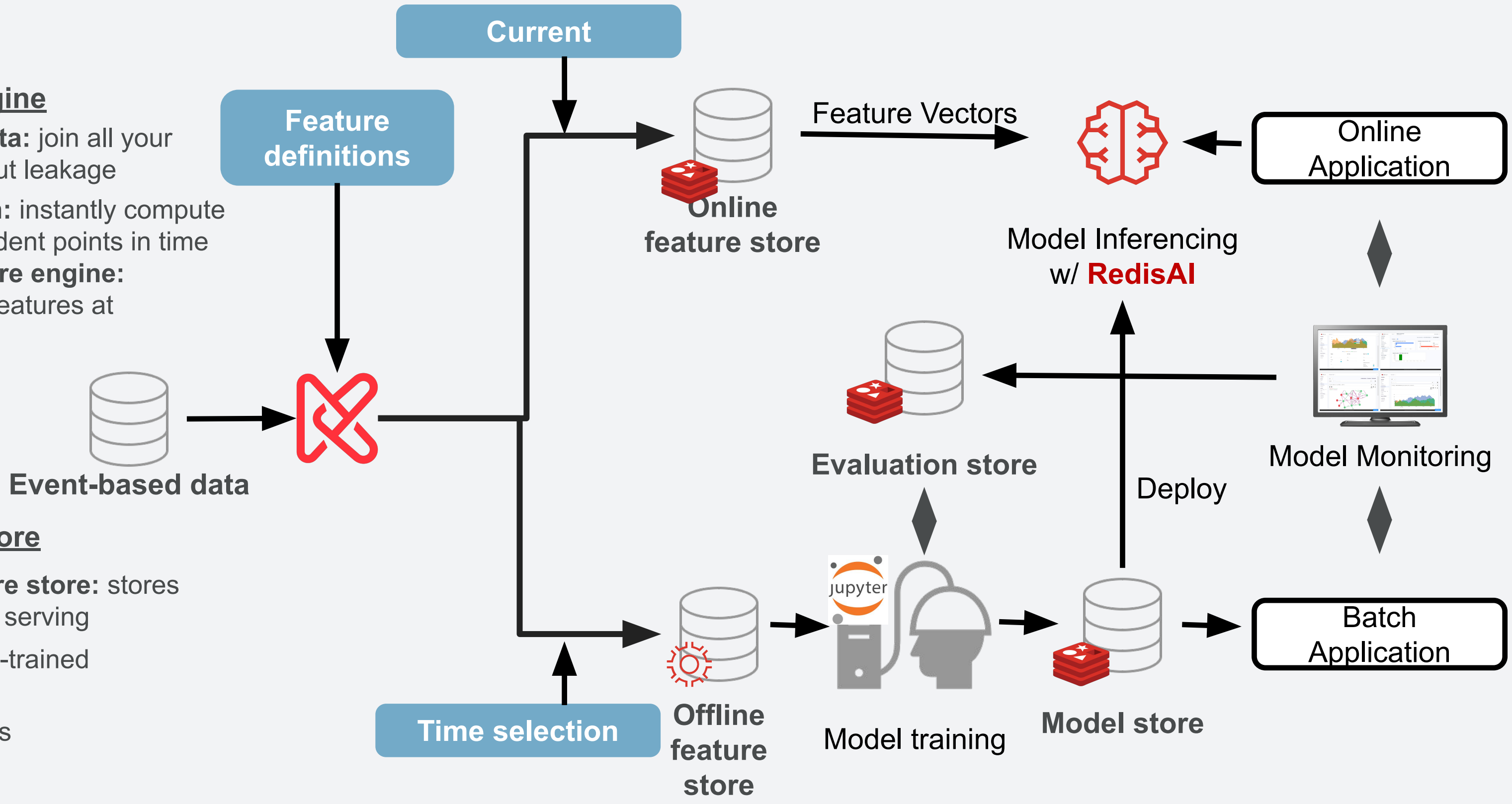
- Over **40% of decision-makers** agree their architectures are not good enough to meet the demands of ML
- The high demand for real-time model inferencing (using ML models in production) expose major challenges with **accuracy**, **latency**, and **reliability** in current architectures
- Running ML model **inferencing in-database** where data is stored solves some of these critical challenges

Forrester Consulting, 2021

Core ML + AI Computing and Serving for Production Stage

Kaskada as a feature engine

- **Connect directly to data:** join all your event-based data without leakage
- **Flexible time selection:** instantly compute at arbitrary, data-dependent points in time
- **Offline & Online Feature engine:** compute and maintain features at relevant points in time



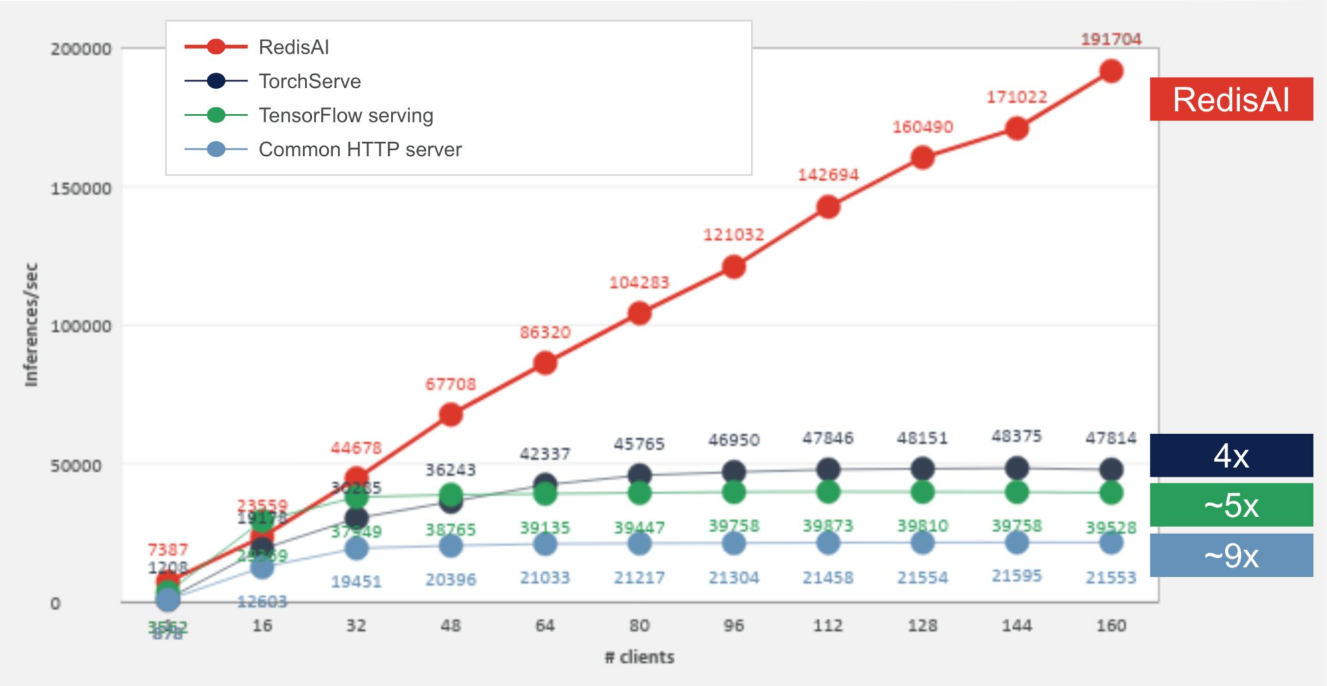
Redis as primary data store

- **Online & Offline feature store:** stores features for low-latency serving
- **Model store:** stores pre-trained models (binaries)
- **Evaluation store:** stores response of the models

redis KASKADA

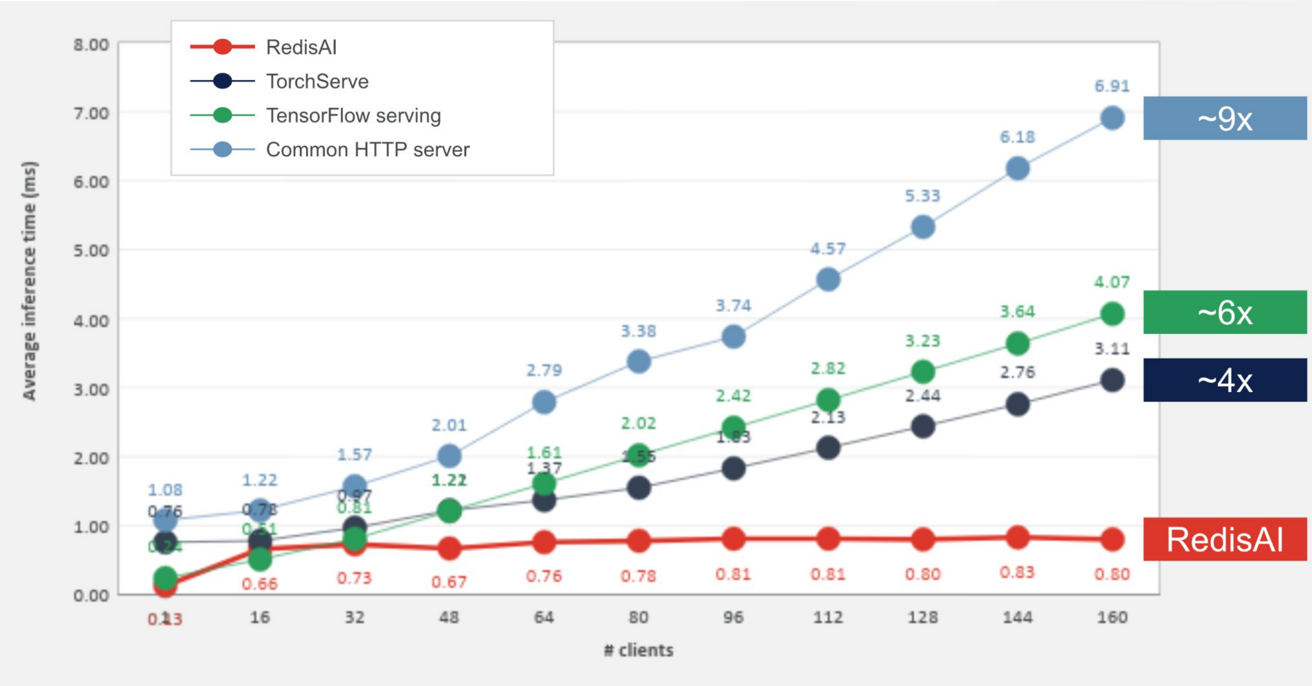
Feature Stores for ML

High throughput required for scoring



Feature Stores for ML

Low latency required for serving



redis KASKADA

Feature Stores for ML

Demo

Summary

Creating models from event based data

- Compute features directly from event based data in order
- Enable iteration by exposing time selection + feature definitions in the feature engineering process
- Join values between different entities at precise times historically to prevent leakage
- Instantly compute values at arbitrary data dependent points in time — discrete and continuous

Operating models from event based data

- Eliminate data discrepancies in production via shared feature definitions
- Low latency applications need a feature store to run model inference in database where the data is stored
- Address real-time throughput needs with a high-throughput feature store

redis KASKADA

Feature Stores for ML

Thank you!

Do you have any questions?

Davor Bonaci
CEO @ Kaskada

@BonaciDavor

Dr. Charna Parkey
VP, Product @ Kaskada

@CharnaParkey

Taimur Rashid
Chief Business Development Officer @ Redis

@taimurrashid