Yelp Ad Targeting at Scale with Apache Spark

Joseph Malicki, Inaz Alaei-Novin

Background - Yelp

Ad Targeting Intro

Model Training

Tools

Deployment to Production

Wrap-up

About us

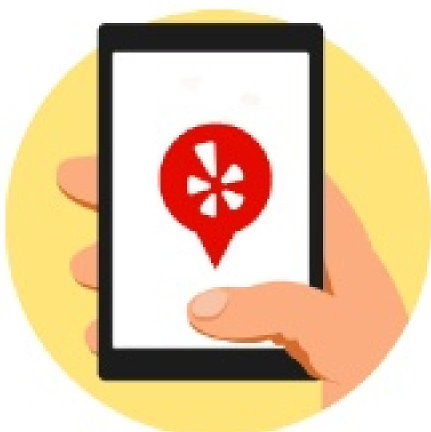
- Joseph Malicki, Inaz Alaei-Novin
- Data mining engineers at yelp
- Ad delivery team

Yelp's Mission

Connecting people with great
local businesses.

Yelp Stats

As of Q1 2017



99M
Monthly Unique
Mobile Users

127M
Monthly Unique
Desktop Users

76% of
Searches via
Mobile App

Background - Yelp

Ad Targeting Intro

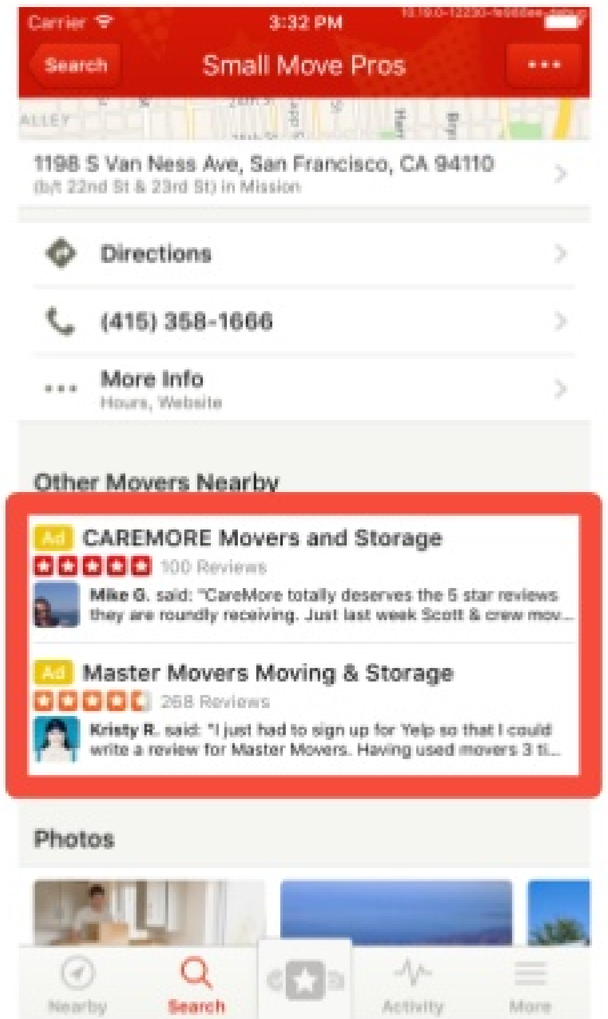
Model Training

Tools

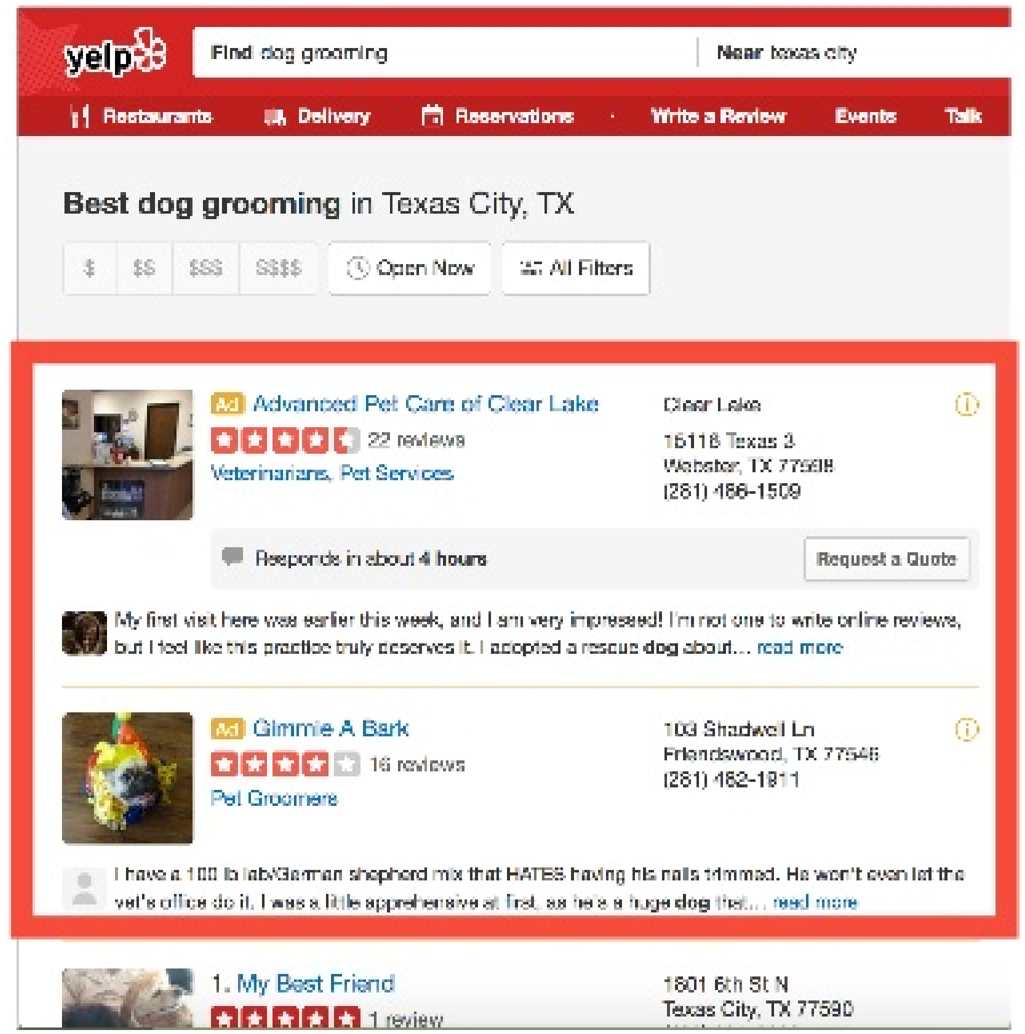
Deployment to Production

Wrap-up

Yelp Ads



Carrier
3:32 PM
Search
Small Move Pros
1198 S Van Ness Ave, San Francisco, CA 94110
(b/t 22nd St & 23rd St) in Mission
Directions
(415) 358-1666
More Info
Hours, Website
Other Movers Nearby
Ad CAREMORE Movers and Storage
100 Reviews
Mike G. said: "CareMore totally deserves the 5 star reviews they are roundly receiving. Just last week Scott & crew mov...
Ad Master Movers Moving & Storage
268 Reviews
Kristy R. said: "I just had to sign up for Yelp so that I could write a review for Master Movers. Having used movers 3 ti...
Photos
Nearby
Search
Activity
More



yelp
Find dog grooming
Near texas city
Restaurants
Delivery
Reservations
Write a Review
Events
Talk
Best dog grooming in Texas City, TX
$ $$ $$$ $$$$
Open Now
All Filters
Ad Advanced Pet Care of Clear Lake
22 reviews
Veterinarians, Pet Services
Clear Lake
15118 Texas 3
Webster, TX 77598
(281) 486-1509
Responds in about 4 hours
Request a Quote
My first visit here was earlier this week, and I am very impressed! I'm not one to write online reviews, but I feel like this practice truly deserves it. I adopted a rescue dog about... read more
Ad Gimmie A Bark
16 reviews
Pet Groomers
103 Shadwell Ln
Friendswood, TX 77546
(281) 482-1811
I have a 100 lb lab/German shepherd mix that HATES having his nails trimmed. He won't even let the vet's office do it. I was a little apprehensive at first, as he's a huge dog that... read more
1. My Best Friend
1 review
1801 6th St N
Texas City, TX 77590

Yelp Ad Targeting

- Majority of Yelp ads are cost-per-click
  - Yelp only gets paid if user clicks on an ad
- Native advertisements
  - Advertisers and content within Yelp platform

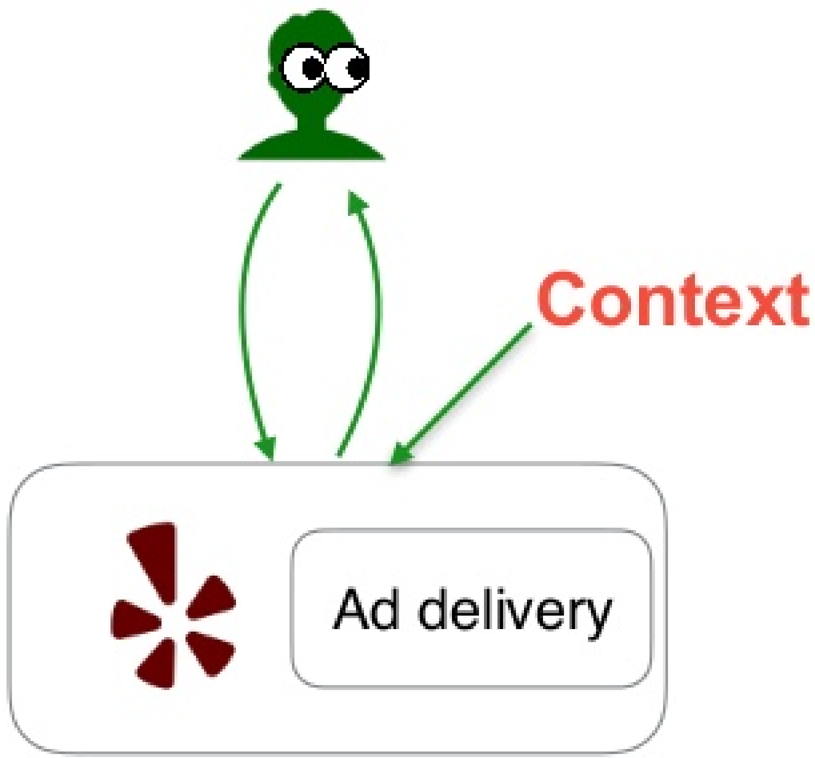
Cost-per-Click Ad Auction

Maximize expected revenue:
1. Order by advertiser bid × predicted click-through-rate (pCTR)
2. Pay second price

Expected[Revenue] = Bid * Expected[CTR]

Because of multiplication, predicted CTR must be well-calibrated, not only well-ordered

Yelp Ad Targeting



- Each request different
- *Context matters*
  - Location
  - Search query
  - User attributes
  - etc.

How to Generalize?

Use machine learning to estimate CTR and show relevant ads

Background - Yelp

Ad Targeting Intro

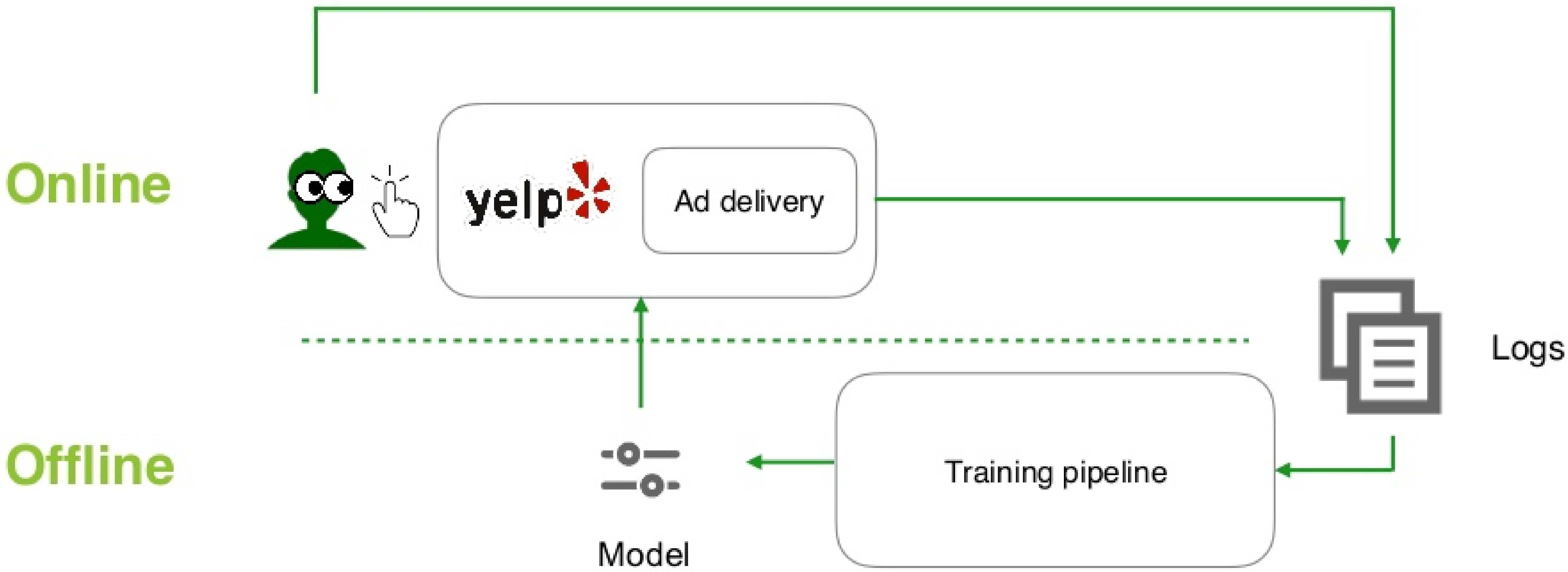
Model Training

Tools

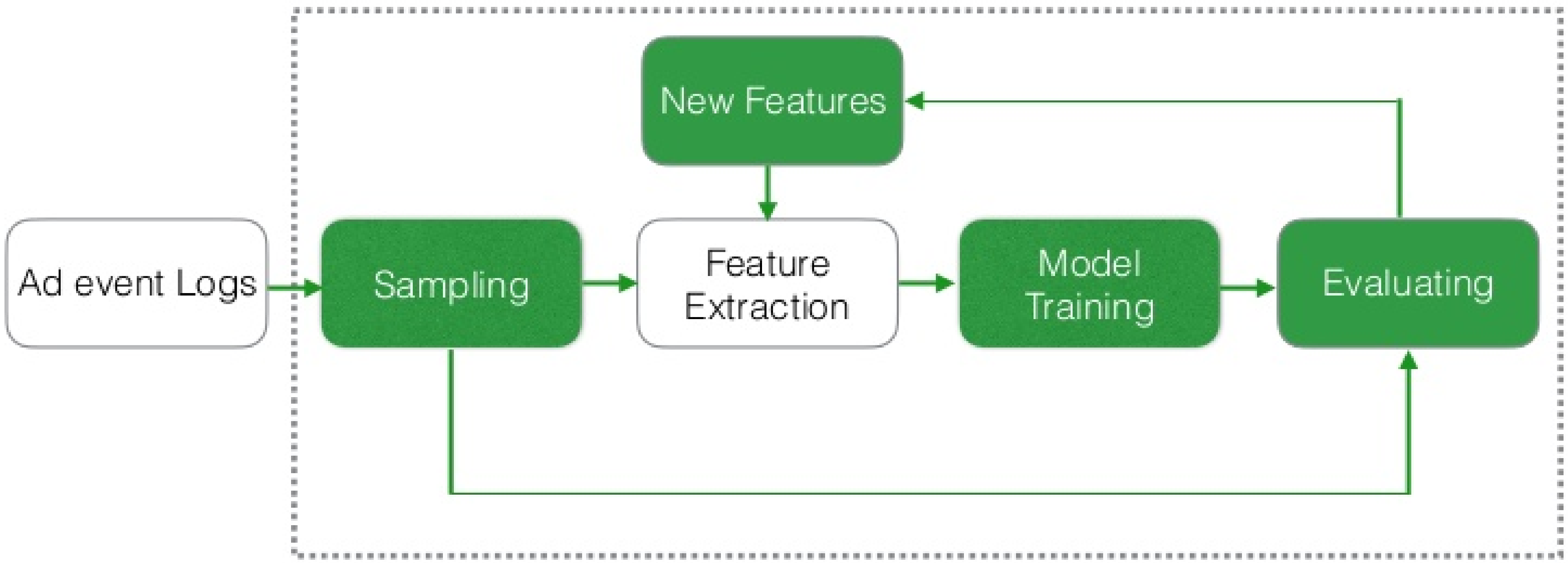
Deployment to Production

Wrap-up

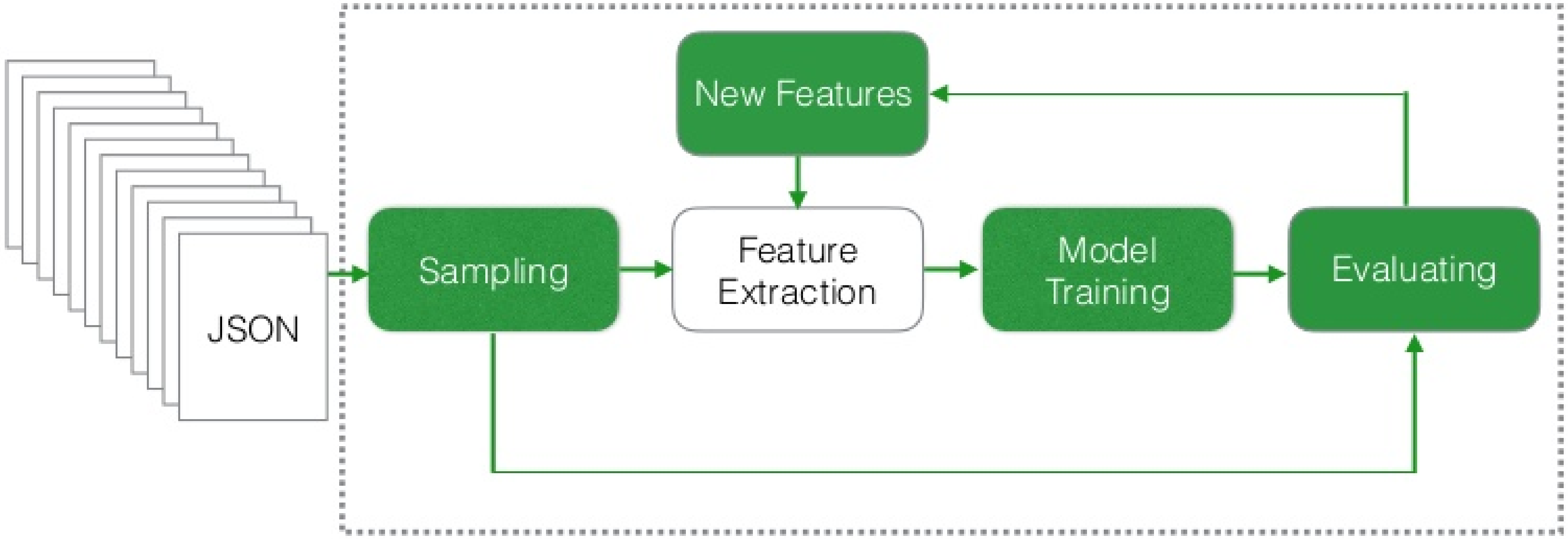
CTR prediction system overview



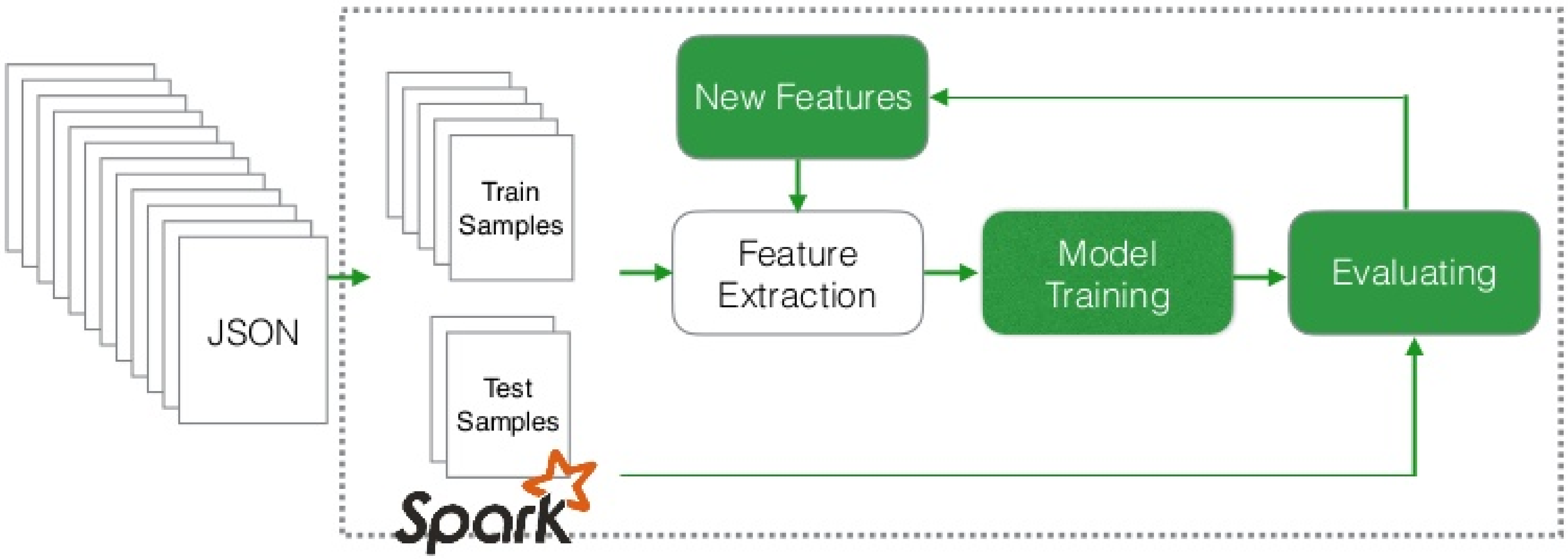
Offline Training at Yelp



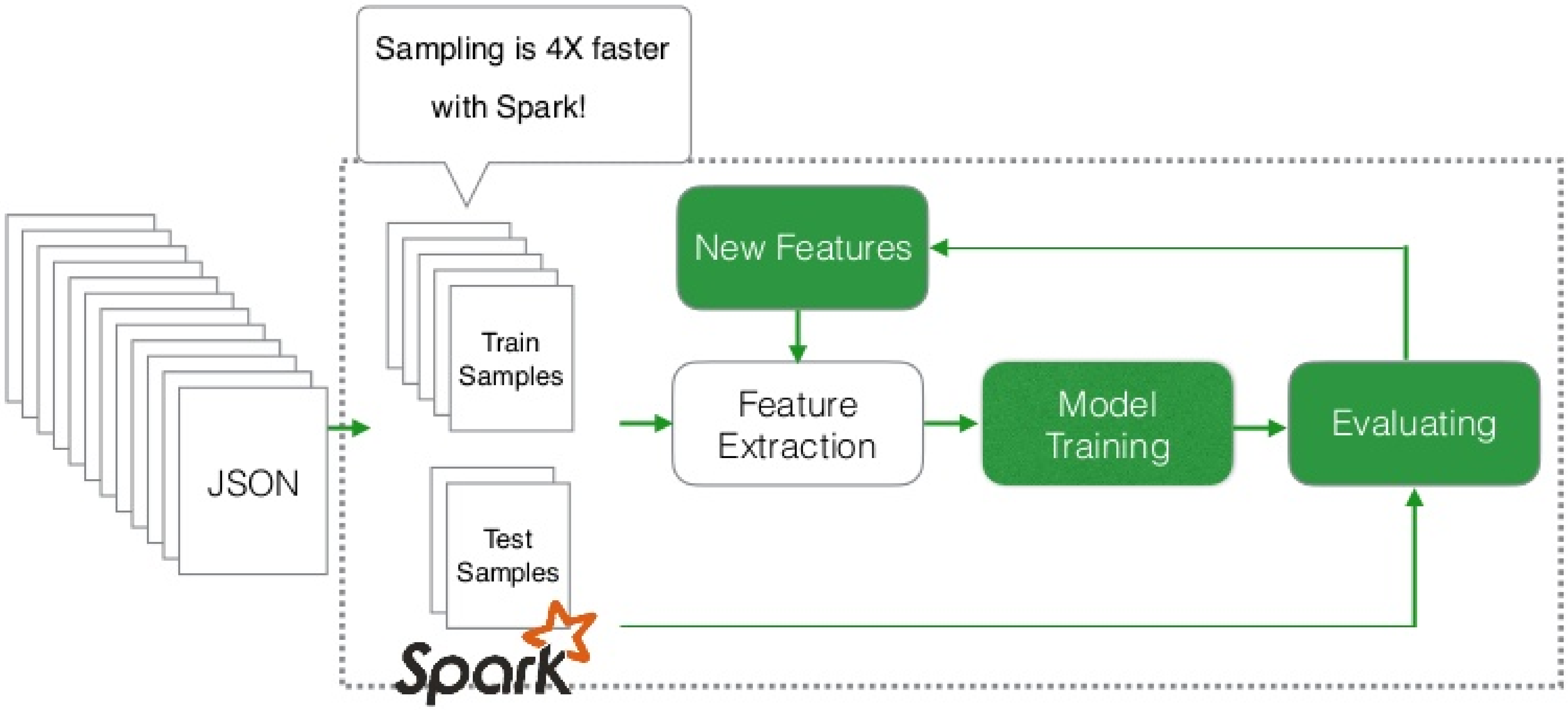
Ad Event Logs



Sampling

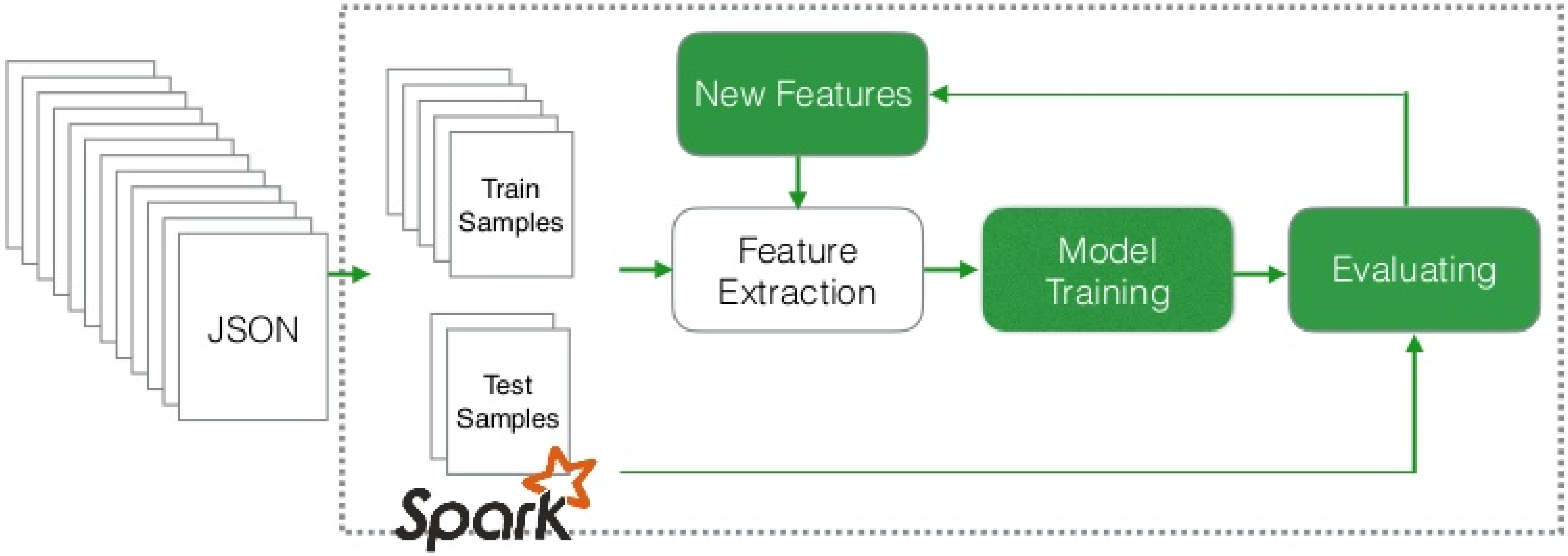

JSON
Train Samples
Test Samples
Spark
New Features
Feature Extraction
Model Training
Evaluating

Sampling



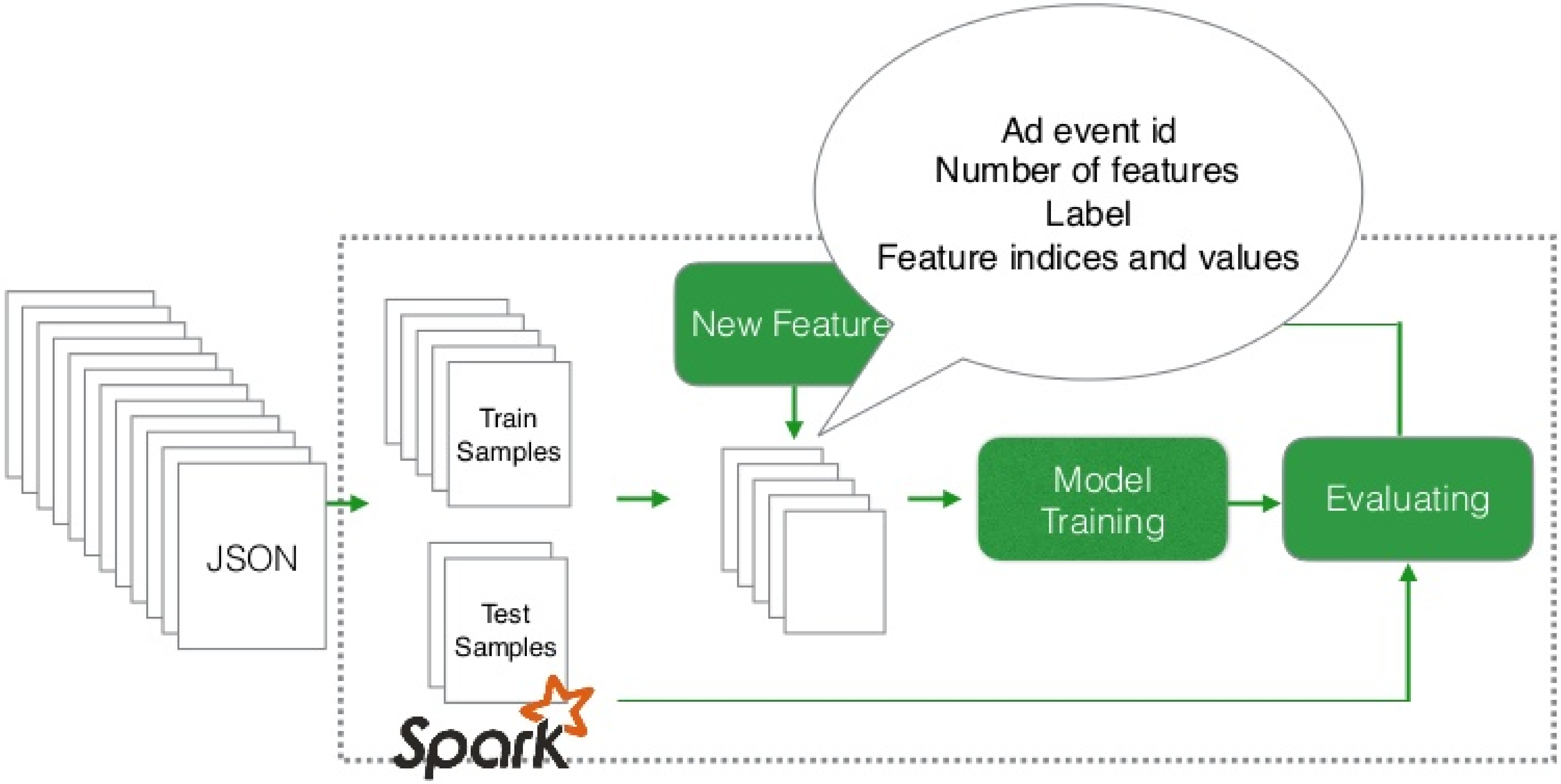
Sampling is 4X faster with Spark!
JSON
Train Samples
Test Samples
Spark
New Features
Feature Extraction
Model Training
Evaluating

Sampling

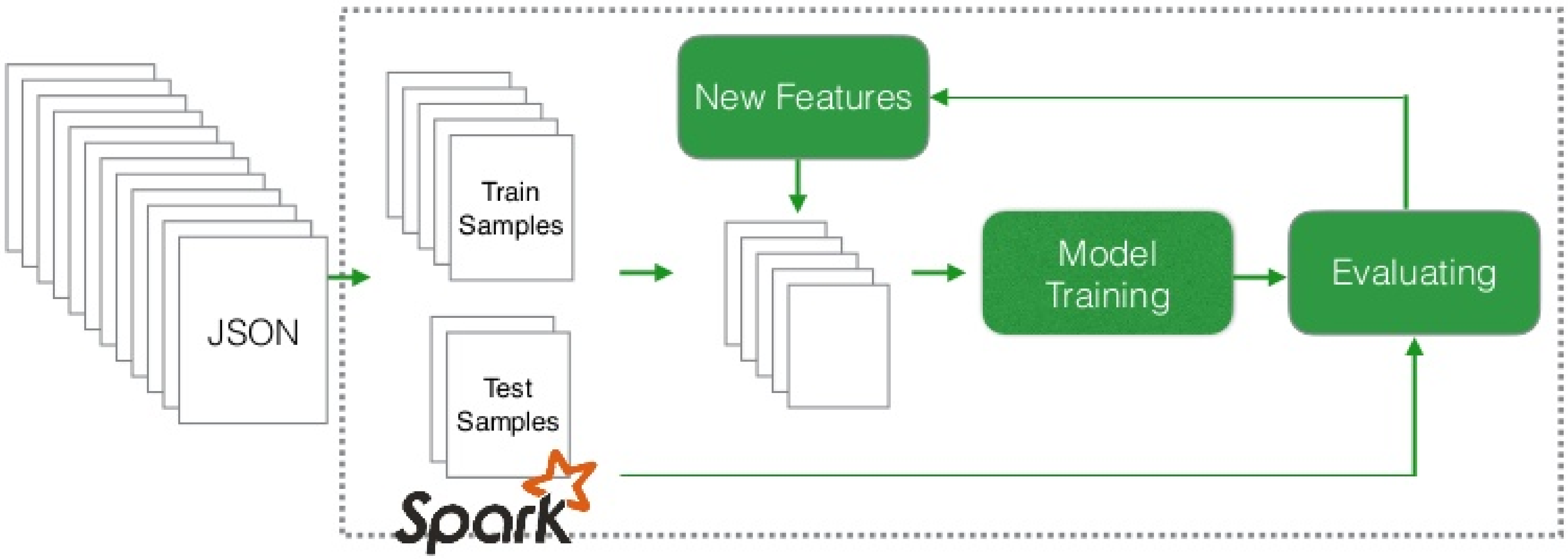


JSON
Train Samples
Test Samples
Spark
New Features
Feature Extraction
Model Training
Evaluating

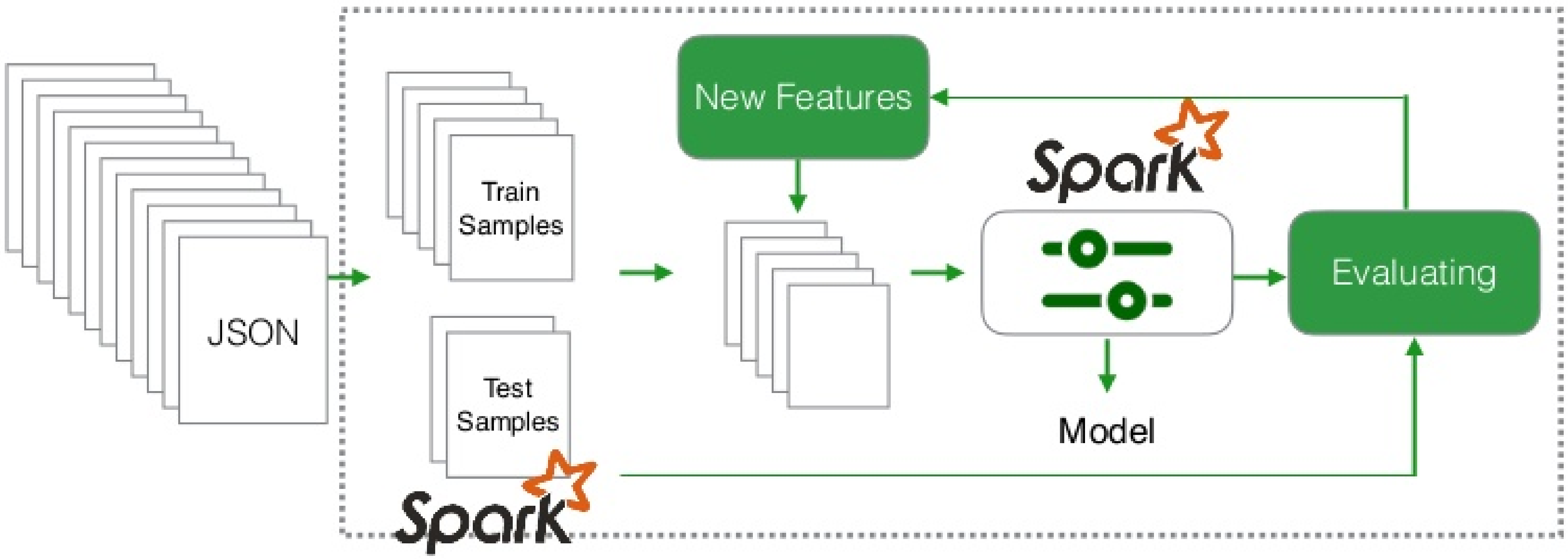
Feature Extraction



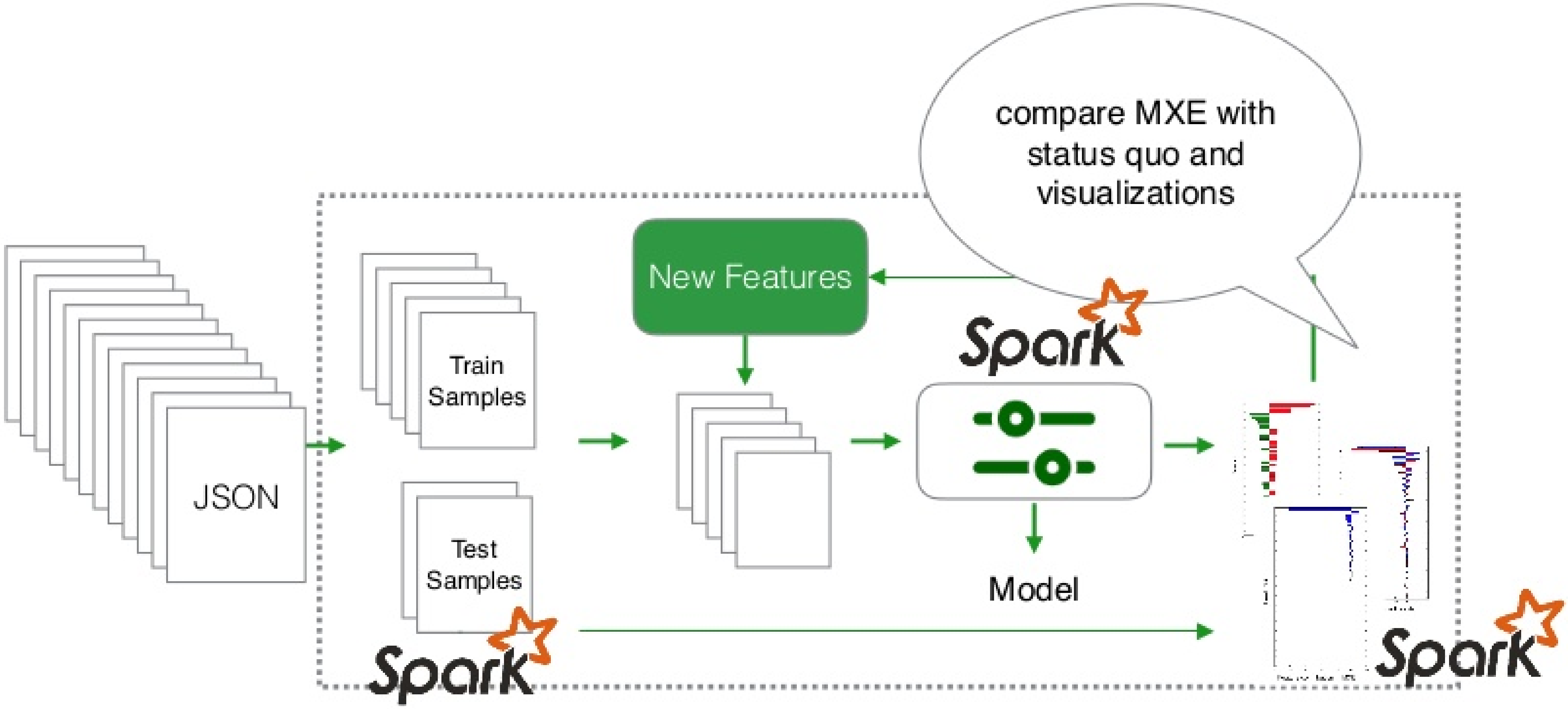
Model Training



Model Training



Evaluation



$$\mathrm{MXE} = -\frac{1}{N}\sum_{n=1}^{N}\left[\, y_n \log \hat{y}_n + (1-y_n)\log(1-\hat{y}_n)\,\right]$$

Feature contribution in a Model

Feature contribution in model

Feature contribution (i) = $\sigma_i \omega_i$

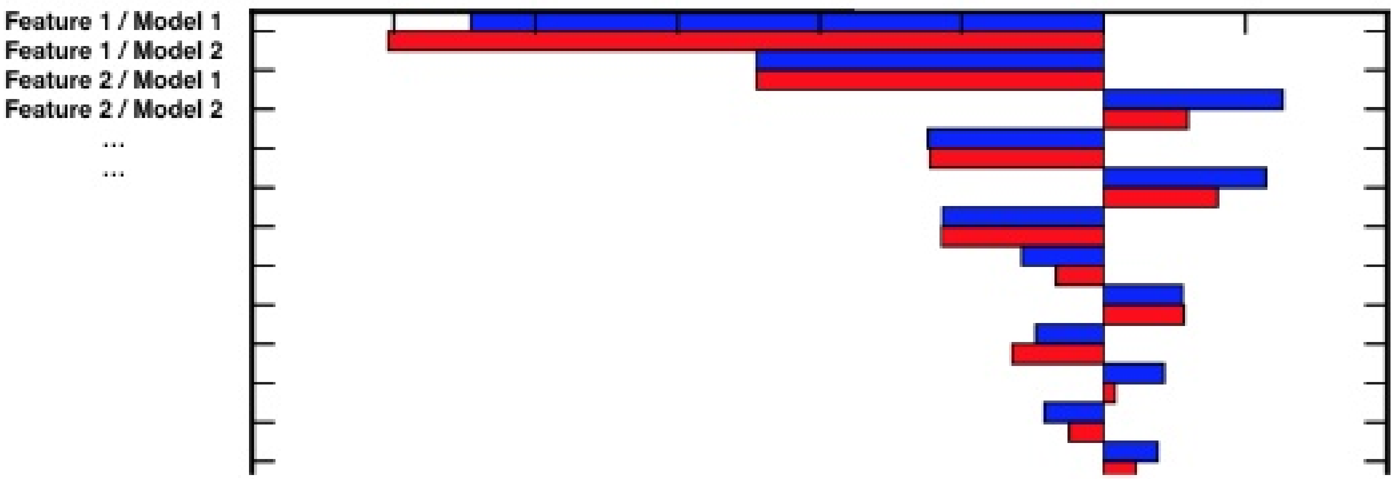
Standard deviation * model coefficient

Compare Feature Importance in Multiple Models

Feature contribution (i) = $\mu_i \omega_i$

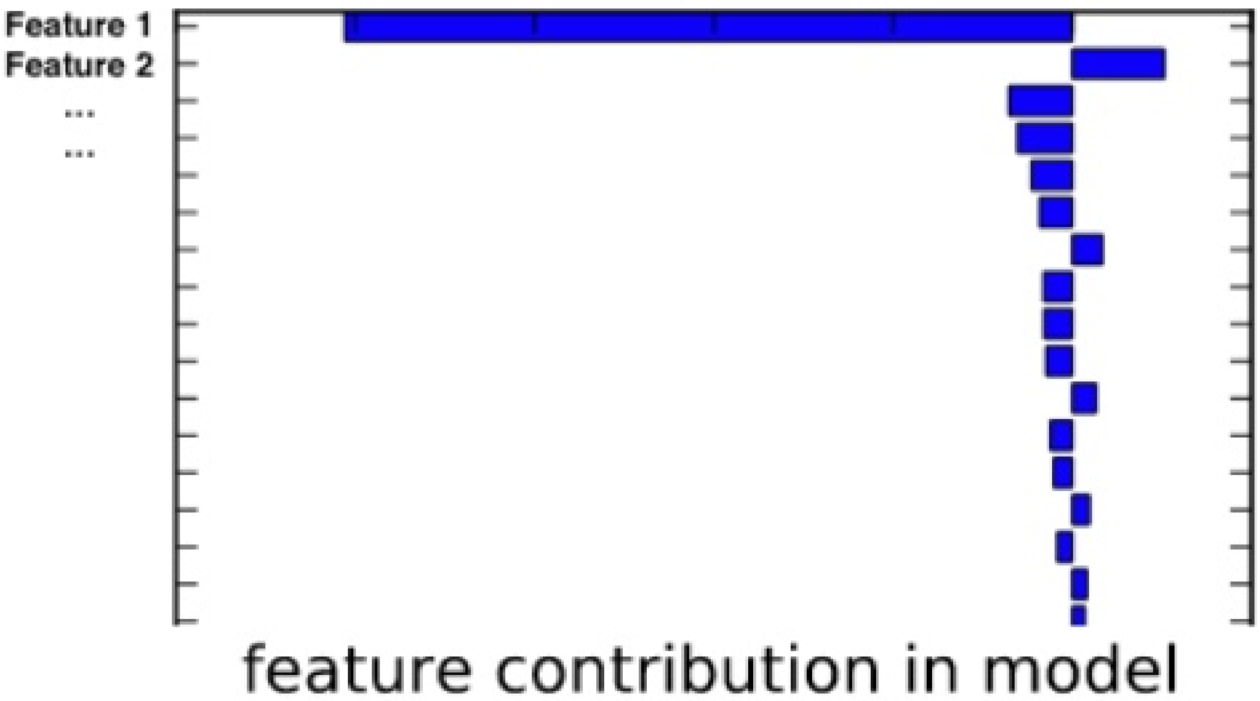
Feature mean * model coefficient

Use colStats from pyspark.mllib.stat.Statistics to compute column summary statistics



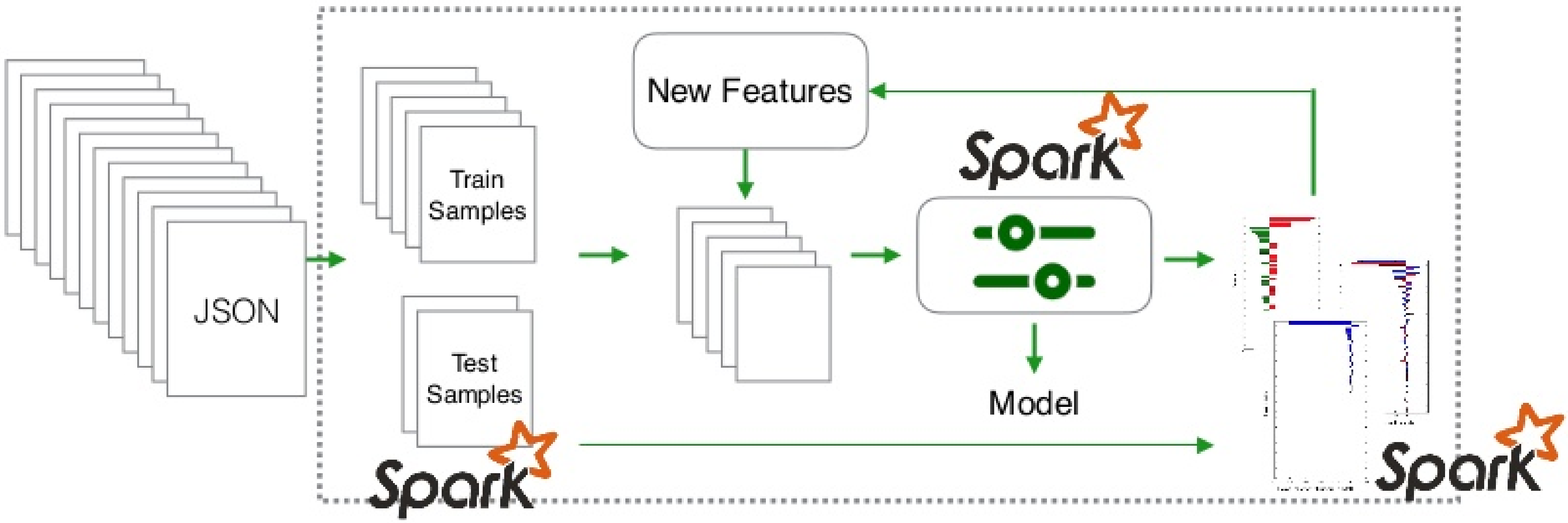
Feature importance in models

Compare Feature Contributions in Models

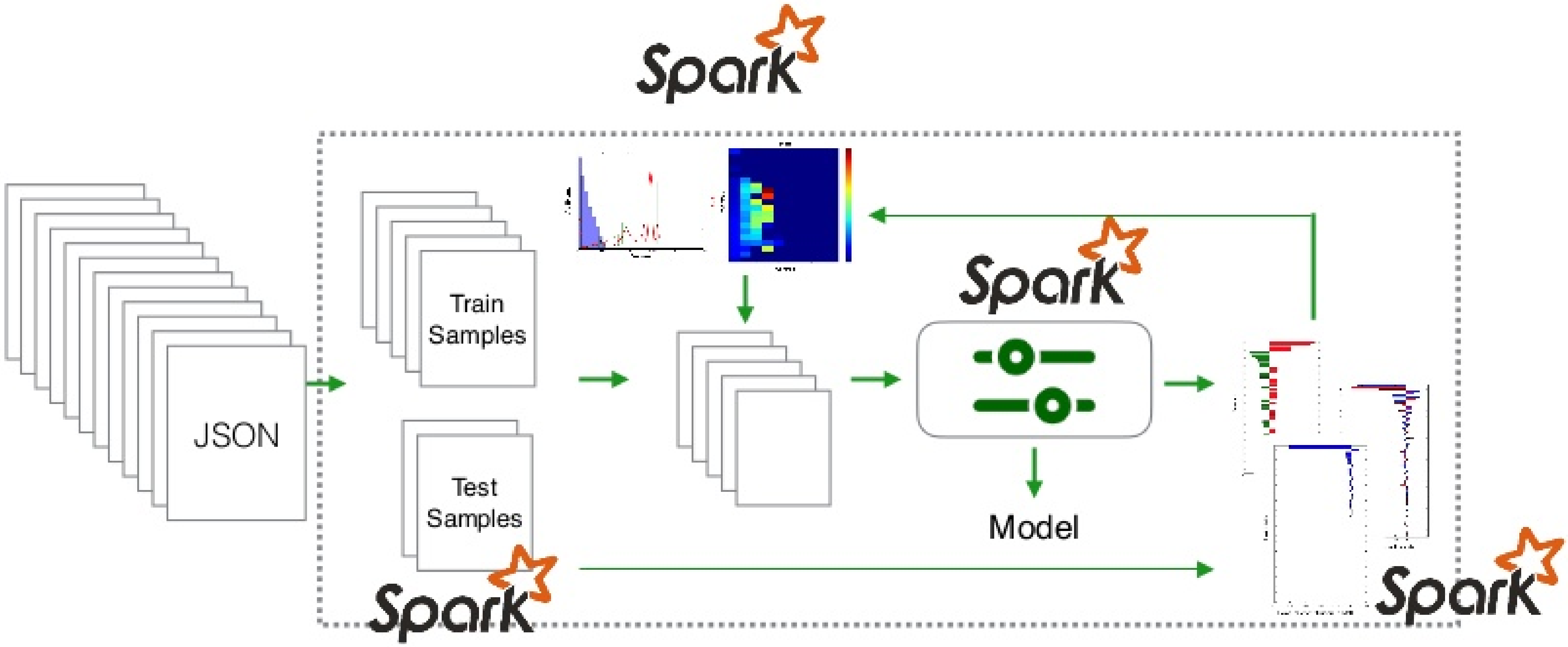


Compare feature contribution in 2 models:
- How much would status quo MXE change if we change the coefficient of one feature from status quo to challenger?

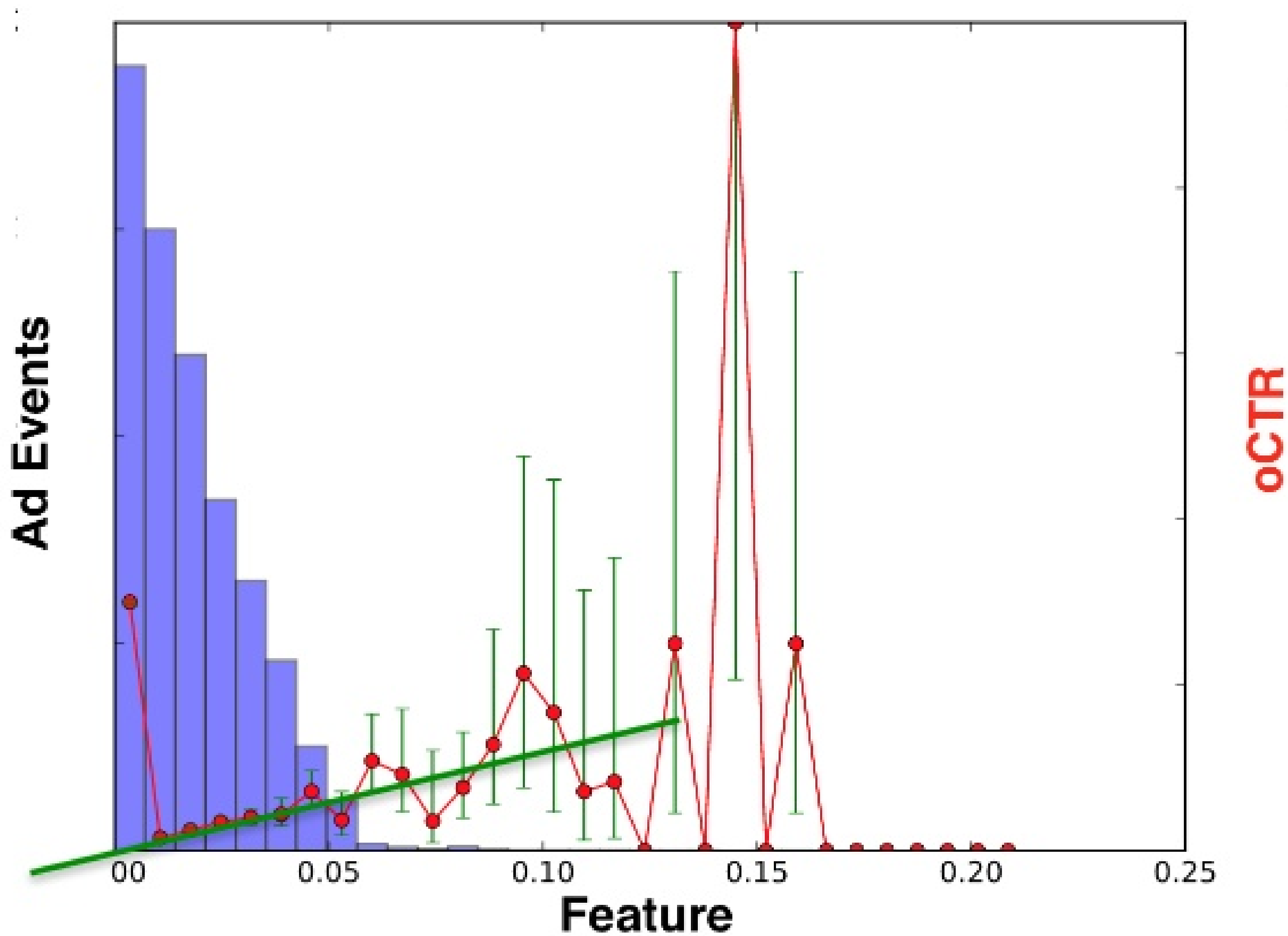
New Features



New Features

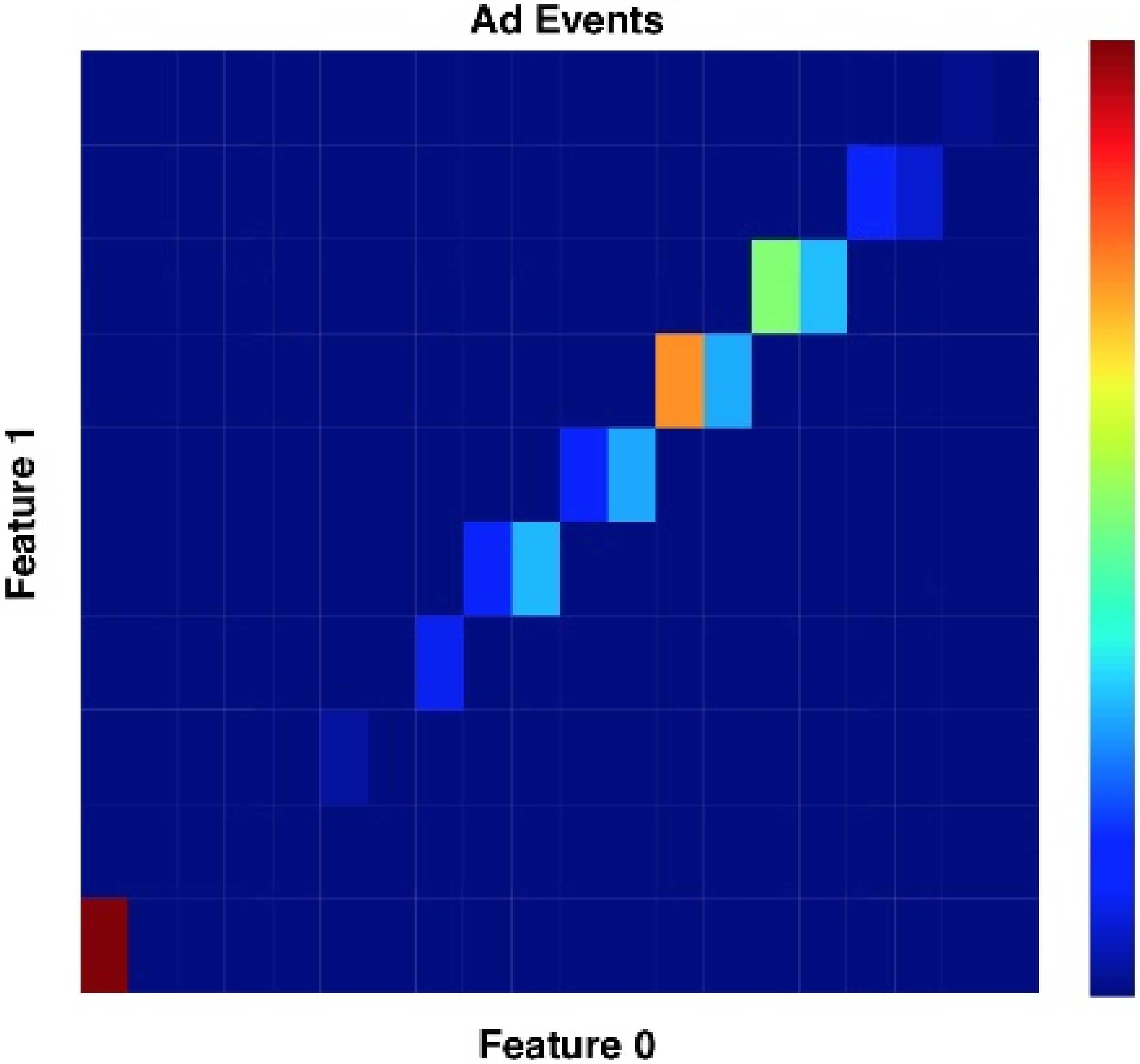


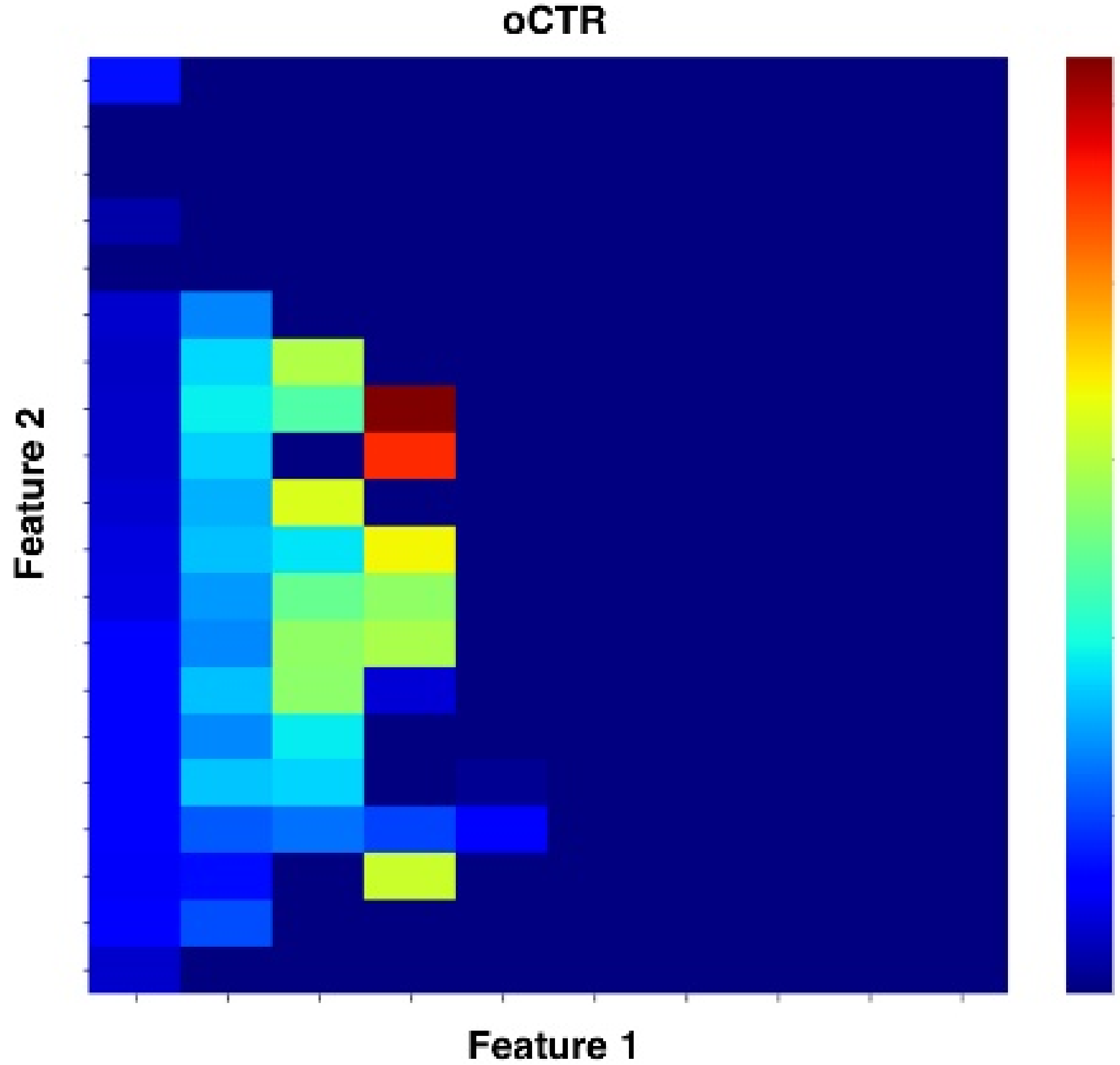
Visualizations



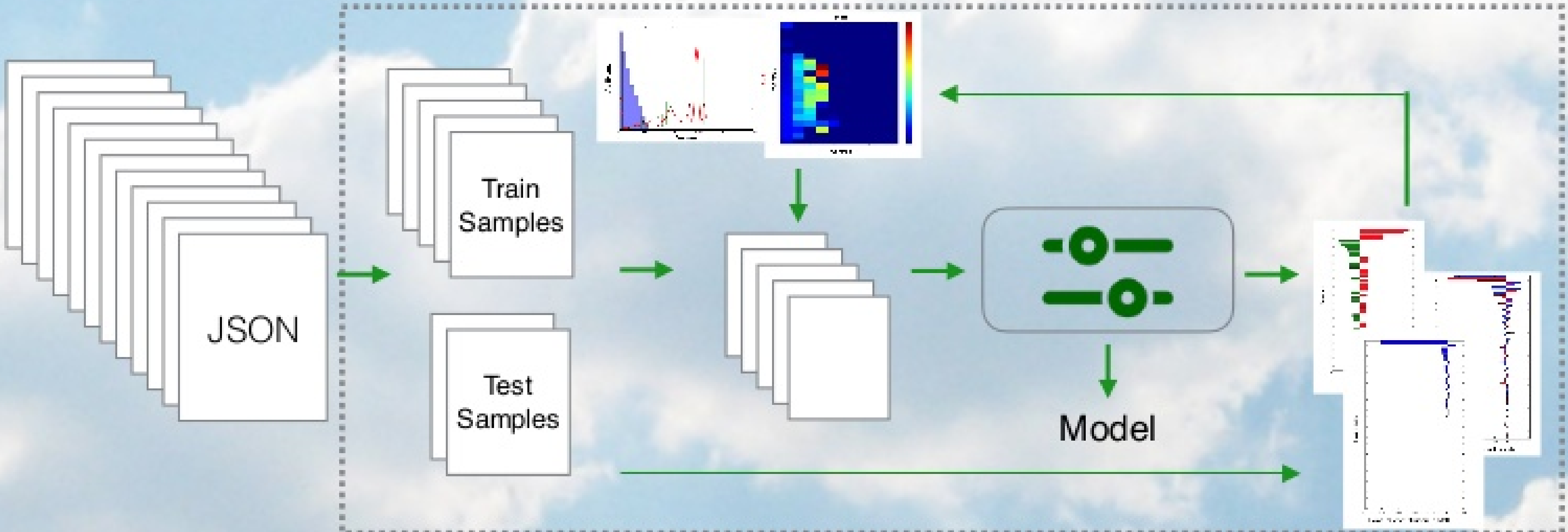
Use RDD's Histogram method and some RDD mappings to generate the plots

Visualizations





Training Pipeline



Background - Yelp

Ad Targeting Intro

Model Training

Tools

Deployment to Production

Wrap-up

Spark related tools

- Zeppelin Notebook
- mrjob

Zeppelin Notebook

- Web-based notebook
- Interactive data analytics
- Supports multiple languages
- Supports Spark
- At Yelp we use it for:
  - Ad-hoc analysis
  - Testing new training algorithms
  - Debugging

mrjob

- One of Yelp's contribution to open source!
- Lets you Write multi-step MapReduce jobs in Python
- Test on your local machine
- Run on a Hadoop cluster
- Run in the cloud using EMR
- Run in the cloud using Google Cloud Dataproc
- Easily run Spark jobs on EMR or your own Hadoop cluster

Background - Yelp

Ad Targeting Intro

Model Training

Tools

Deployment to Production

Wrap-up

Production concerns

Offline Batch

- Overnight or developer-initiated jobs
- Millions to billions of datapoints
- Batch-oriented (hours)
- Apache Spark

Online Ad Serving

- User hits button on app, needs quick response
- Smaller number of locally and contextually relevant candidates
- Real-time (milliseconds)
- Java servlet

Shared code (libraries)

SPARK SUMMIT 2017

Monitoring

- If CTR prediction model stops being accurate, could lead to loss of revenue
- How do we know models are working properly?
- Need to check model predictions are accurate over time

Monitoring

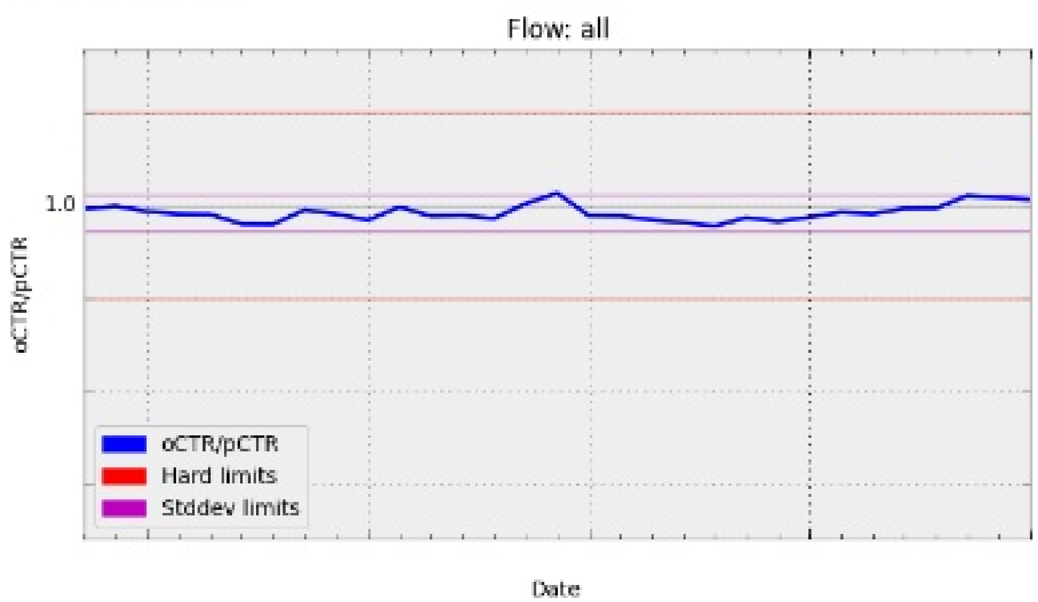
- Large batch jobs check actual user ad click-through-rate against predicted CTR
- Model accuracy far more sensitive than overall metrics: traffic mix is accounted for
- Spark streaming allows real-time alerts

  - A practical approach to building a streaming processing pipeline for an online advertising platform - Spark Summit 2017

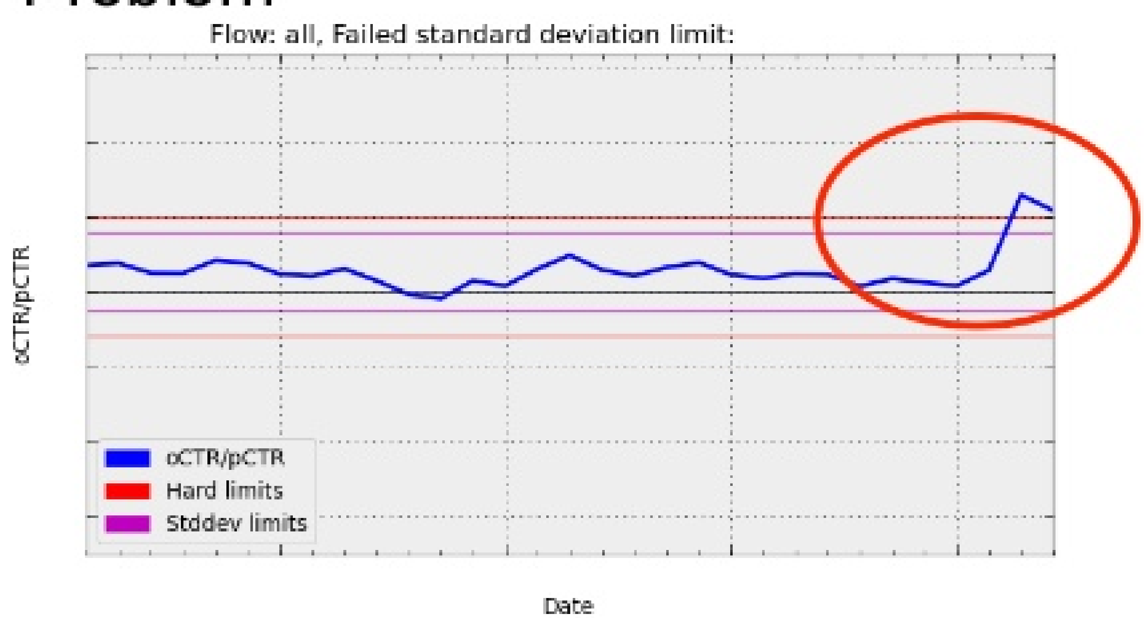
Monitoring - Examples

- Misspelled header in API call refactor
- Change in HTTPS caching behavior affects CTR

Normal



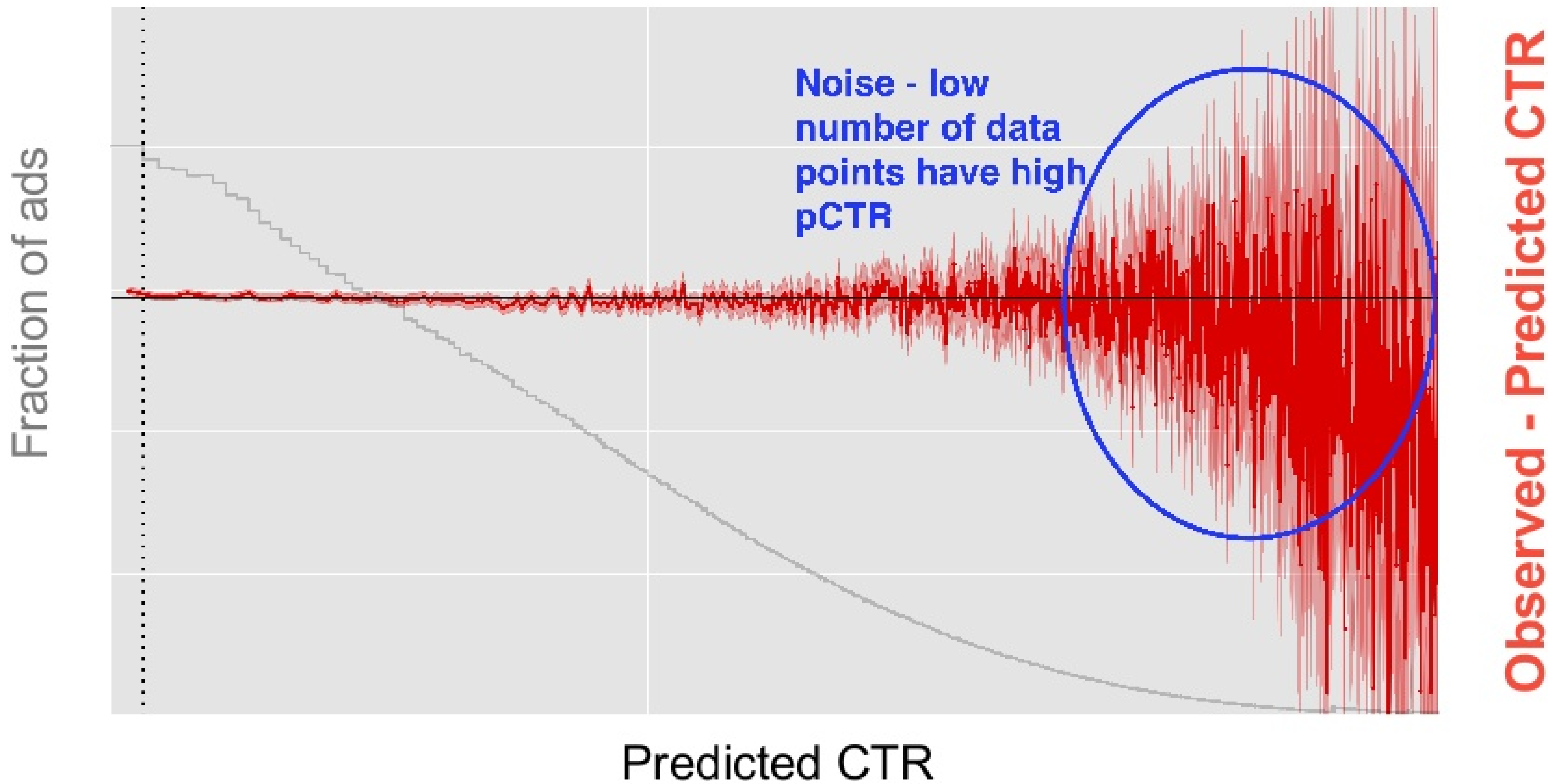
Problem



Monitoring - Calibration Plot

- Recall ad auction orders by advertiser bid × predicted click-through-rate (pCTR)
- Because of multiplication, predicted probabilities need to be well-calibrated
- Goal: $P\left(clicked \mid C\hat{T}R = y\right) = y$

Monitoring - Calibration Plot



Monitoring - Calibration Plot

- Logistic regression loss is a *proper scoring rule*
  - Generates models that are well-calibrated on average
- Feature engineering problems can cause poor calibration
- Probability distribution drifting over time will cause loss of calibration
  - e.g. changes to user interface affecting behavior

Background - Yelp

Ad Targeting Intro

Model Training

Tools and Visualizations

Deployment to Production

Wrap-up

Spark at Yelp

- Spark increasingly used throughout Yelp
  - Streaming
  - Iteration
  - Easy specification of job flows
- Want to work with Spark? We're hiring - stop by Yelp booth in exhibition area, until 4:30pm


We're Hiring!
www.yelp.com/careers/
EAT24

fb.com/YelpEngineers

@YelpEngineering

engineeringblog.yelp.com

github.com/yelp

Thank You.

Questions?