

# LTR at GetYourGuide Marketplace

A Journey through our experience

Ashraf Aaref and Felipe Besson

June 13th 2018

**MICES 2018**

MIX-CAMP E-COMMERCE SEARCH

# Who are we?

We work for the search team at GYG

- Ashraf
  - Software Engineer
- Felipe
  - Data Engineer



# Agenda?

- What is GetYourGuide and our challenges?
- V1: Our first try to apply LTR
- Lesson learned
- Next step, V2?
- Questions

# What is GetYourGuide?

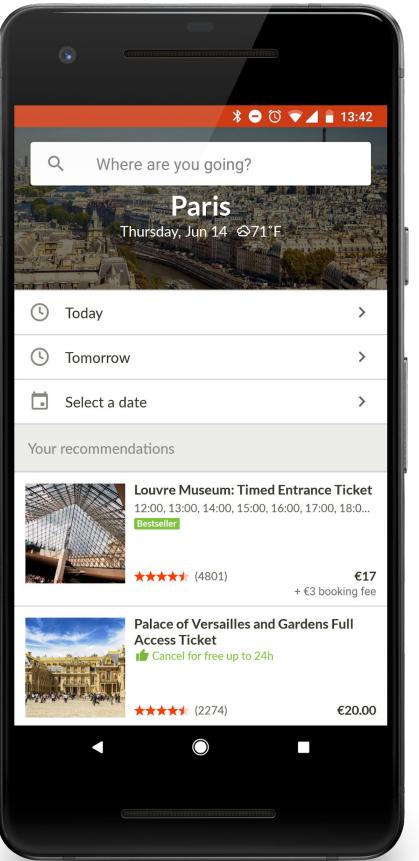
GetYourGuide is a marketplace for activities, such as guided tours, ticketed attractions, airport transfers, different experiences, and more...

- +33K Activities
- +20 Languages
- +7K destinations
- +400 Employees



# Full-text Search

- Location driven
- Discovery



Get Your Guide

English ▾ EUR (€) ▾ Cart Help

Book tickets for top attractions around the world

Where are you going?  from  to

GetYourGuide gives you:

- The best selection More than 34,520 things to do
- The lowest prices We guarantee it!

### Top attractions

Eiffel Tower

Vatican Museums

Louvre Museum

Burj Khalifa

Empire State Building

Colosseum

# Location pages (LPs)

- Location driven
- Dates are **very** important
- **High-intent customers**
- **Paid traffic**



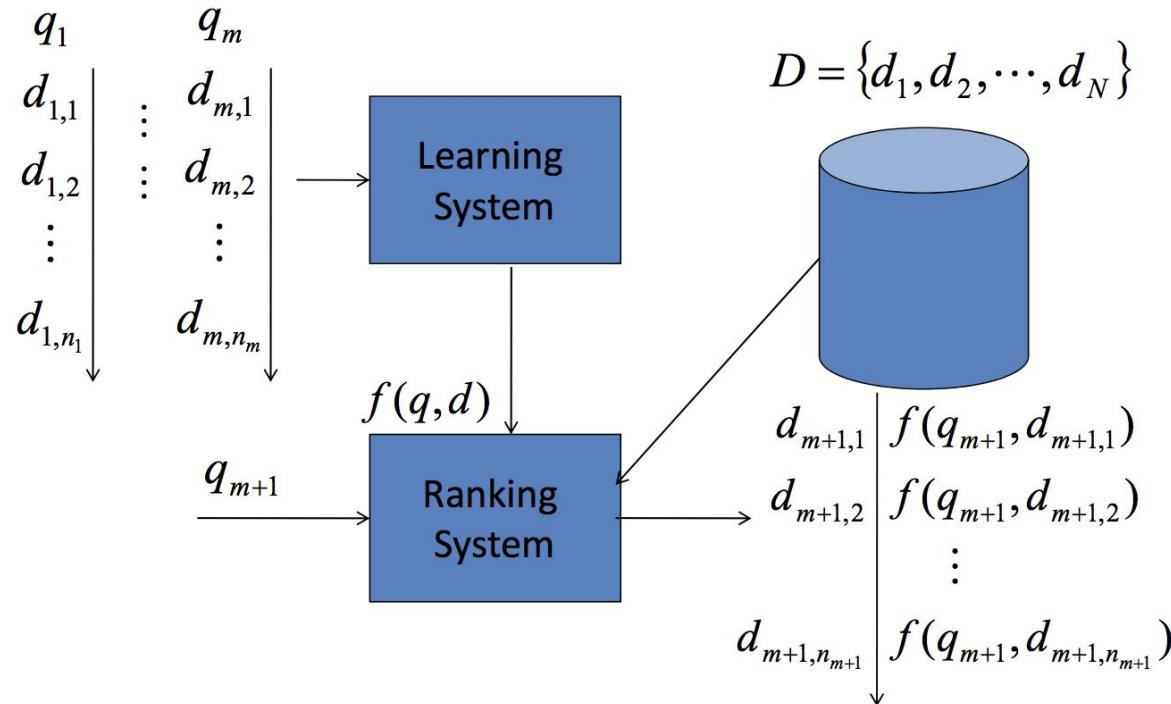
GetYourGuide

The screenshot shows the GetYourGuide website for Paris. At the top, there's a red banner with the text "Get Your Guide". The main header is "Paris" with the subtext "Must-see sights and things to do". Below the header is a search bar with fields for "from" and "to" dates, and a "Check availability" button. The background features a large image of the Paris skyline with the Eiffel Tower. Below the search bar, there are three promotional boxes: "GetYourGuide gives you:" (The best selection: More than 34,520 things to do), "The lowest prices: We guarantee it!", and "Fast & easy booking: Book online to lock in your tickets". The main content area is titled "Top attractions in Paris" and lists several options with images, descriptions, and ratings. The attractions include: Eiffel Tower Walk to the 2nd Floor & Optional Summit Access (From €32, 402 reviews), Louvre Museum: Timed Entrance Ticket (From €17 (+ €3 booking fee), 4681 reviews), Paris Illuminations 1-Hour Cruise (From €15, 2419 reviews), Palace of Versailles and Gardens Full Access Ticket (From €20, 2233 reviews), Paris: Skip-the-Line Arc de Triomphe Rooftop Tickets, Paris: 1, 2, or 3-Day Hop-on Hop-off Ticket, Eiffel Tower: Second Floor Priority Access, and Disneyland Paris 1-Day Ticket.

# Problems with LP Ranking

- Focus on business metrics
- Customer intentions (search keywords)
  - "Eiffel Tower ticket" = "Eiffel Tower restaurant"
- Difficult to introduce new and diverse products
- **We needed to learn how to rank activities in LPs!**

# Let the machine do it for you! (LTR)



Extracted from ACML 2009 Tutorial Nov. 2, 2009 Nanjing



# First iteration (V1)

## Scope and decisions

# Learning to Rank (LTR) at GYG

Apply Machine Learning to introduce **relevance** factors into our ranking formula

Use our user intention data to have a dynamic LP ranking

# V1 Focus

- Vertical: Points of Interest
  - Ticket, Tour, Museum, Historic site, park, ...
- Only in English (we have 22 languages)
- Location pages have no explicit user query
  - Search Keywords:

**"Statue of Liberty boat tour"**



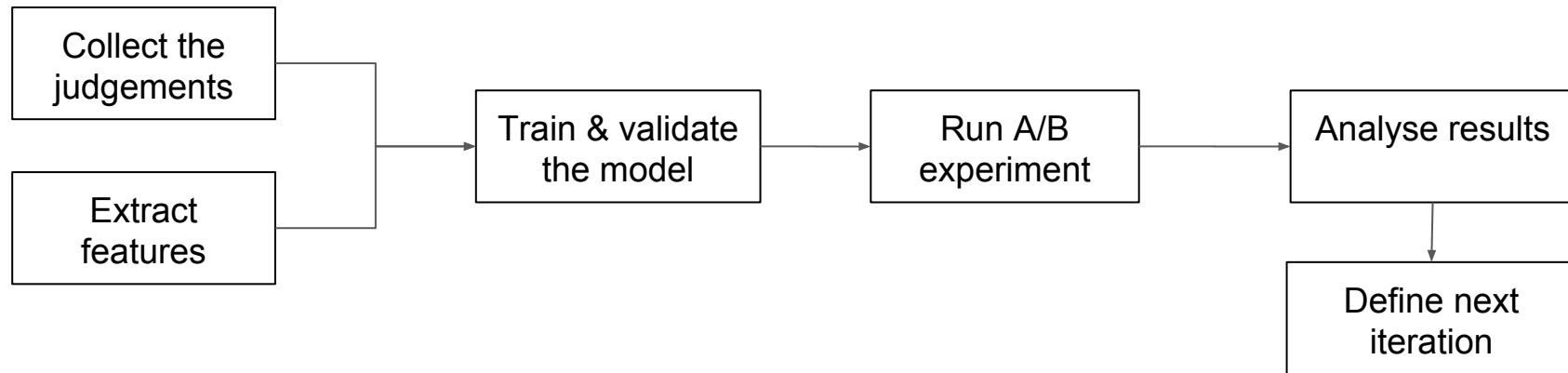
**location**



**intention**

# MVP mindset

Follow the standard steps of a LTR solution



A photograph of a two-lane asphalt road with yellow double lines on the left and a white dashed line on the right. The road curves slightly to the right as it leads into a dark, rocky tunnel. The surrounding environment is a forest of bare, leafless trees and large, light-colored rocks on a hillside. The sky is overcast and grey.

We started the journey!

# Judgement List

q = "Eiffel Tower restaurant"

Document	Judgement
<p> <b>Dinner at 58 Tour Eiffel with Priority Lift Tickets</b> ★★★★★ 701 Reviews Dine at the chic 58 Tour Eiffel located on the first floor of Paris's most iconic monument. Enjoy panoramic views of the city while you feast on your... ⌚ Duration: 2 hours – 2.5 hours From <b>US\$114.14</b></p>	<div style="background-color: #339933; color: white; padding: 5px 10px; border-radius: 10px; text-align: center;">3</div>
<p> <b>Eiffel Lunch, 2nd Floor Ticket and Seine River Cruise</b> ★★★★★ 16 Reviews Experience the second floor's panoramic view on the Eiffel Tower. Enjoy the special combination of lunch at Restaurant 58 on the first floor and... ⌚ Duration: 4 hours From <b>US\$133.56</b></p>	<div style="background-color: #339933; color: white; padding: 5px 10px; border-radius: 10px; text-align: center;">3</div>
<p> <b>Eiffel Tower: Second Floor Priority Access</b> ★★★★★ 5484 Reviews <span style="background-color: red; color: white; padding: 2px 5px; border-radius: 5px;">LIKELY TO SELL OUT</span> Save precious time at the Eiffel Tower with a priority-access ticket to the second level. ⌚ Duration: 2 hours From <b>US\$49.79</b></p>	<div style="background-color: #339933; color: white; padding: 5px 10px; border-radius: 10px; text-align: center;">2</div>
<p> <b>FlyView: Fly over Paris in Virtual Reality</b> ★★★★★ 6 Reviews Put on your jetpack and see Paris from a bird's-eye view. The 15-minute FlyView virtual-reality experience is complete with real flight sensations. ⌚ Duration: 31 minutes From <b>US\$14.57</b></p>	<div style="background-color: #FF9999; color: white; padding: 5px 10px; border-radius: 10px; text-align: center;">1</div>
<p> <b>Tower of London Ticket with Crown Jewels Exhibition</b> ★★★★★ 1495 Reviews <span style="background-color: red; color: white; padding: 2px 5px; border-radius: 5px;">LIKELY TO SELL OUT</span> Visit the Tower of London and see the famous building that has served as a fortress, palace, and prison. ⌚ Duration: 60 minutes – 6 hours From <b>US\$37.04</b></p>	<div style="background-color: #FF9999; color: white; padding: 5px 10px; border-radius: 10px; text-align: center;">0</div>

# Human labeling judgement list

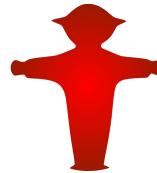
- Judgements were collected from Domain Experts
  - Internal stakeholders of GYG
- Judgement scale
  - **0 - 3**
- ~ 30k judgements
- Pre analysis of current rank
  - NDCG@7 = **0.55**



# Human labeling judgement list



- ✓ Good approach when data is incomplete/inconsistent
- ✓ When what is a relevant result is still unclear
- ✓ No need to normalize queries deeply



- ✗ Relevance is subjective from user to user
- ✗ Hard to scale
- ✗ Crowdsourcing is expensive

# Enriching Judgements with features



# Feature Engineering

Query document	Business metrics	Document
<ul style="list-style-type: none"><li>• BM25 of single text fields</li><li>• Multi-match combinations</li></ul>	<ul style="list-style-type: none"><li>• Raw metrics: clicks, bookings, impressions</li><li>• Rates: CTR, CR</li></ul>	<ul style="list-style-type: none"><li>• Activity attributes: price, duration, # reviews</li></ul>

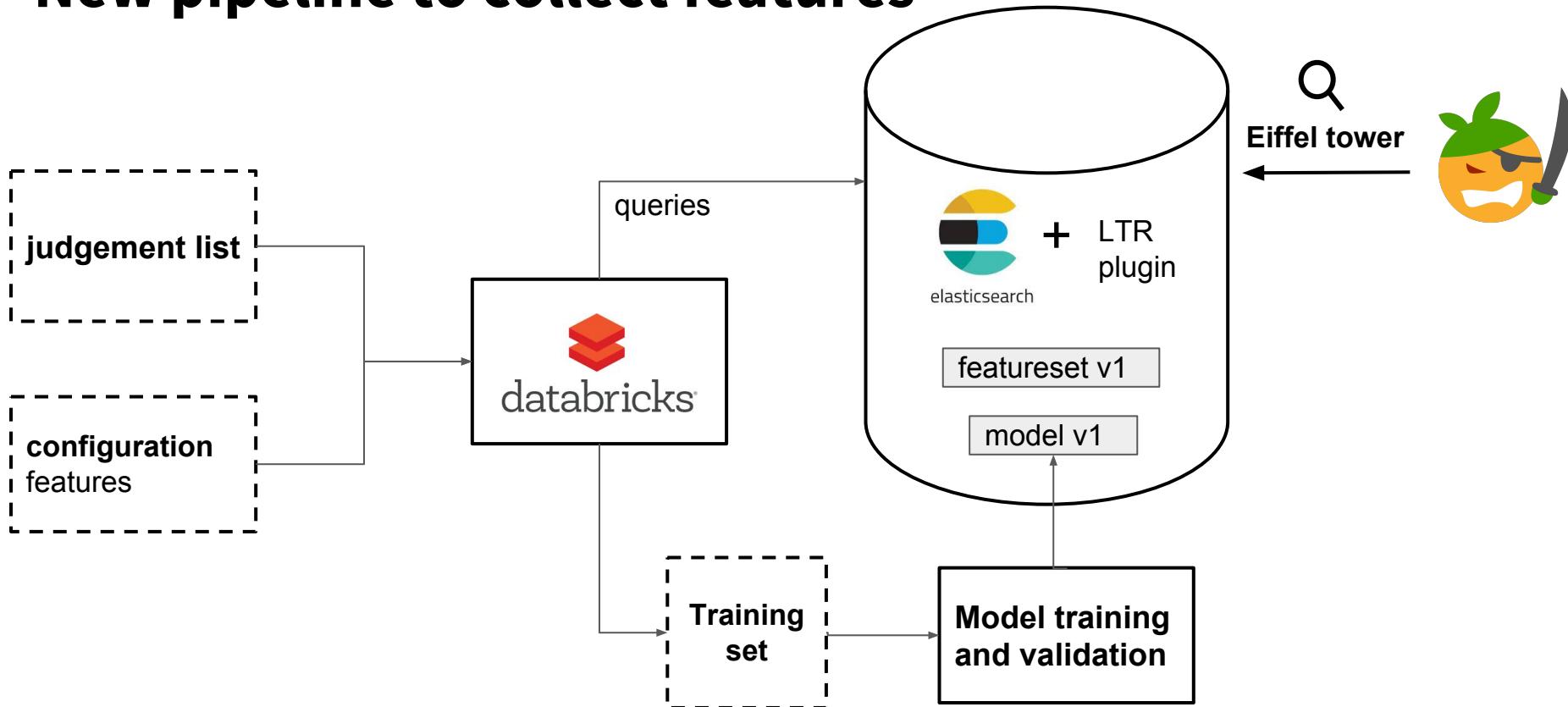


# How to collect these features ?

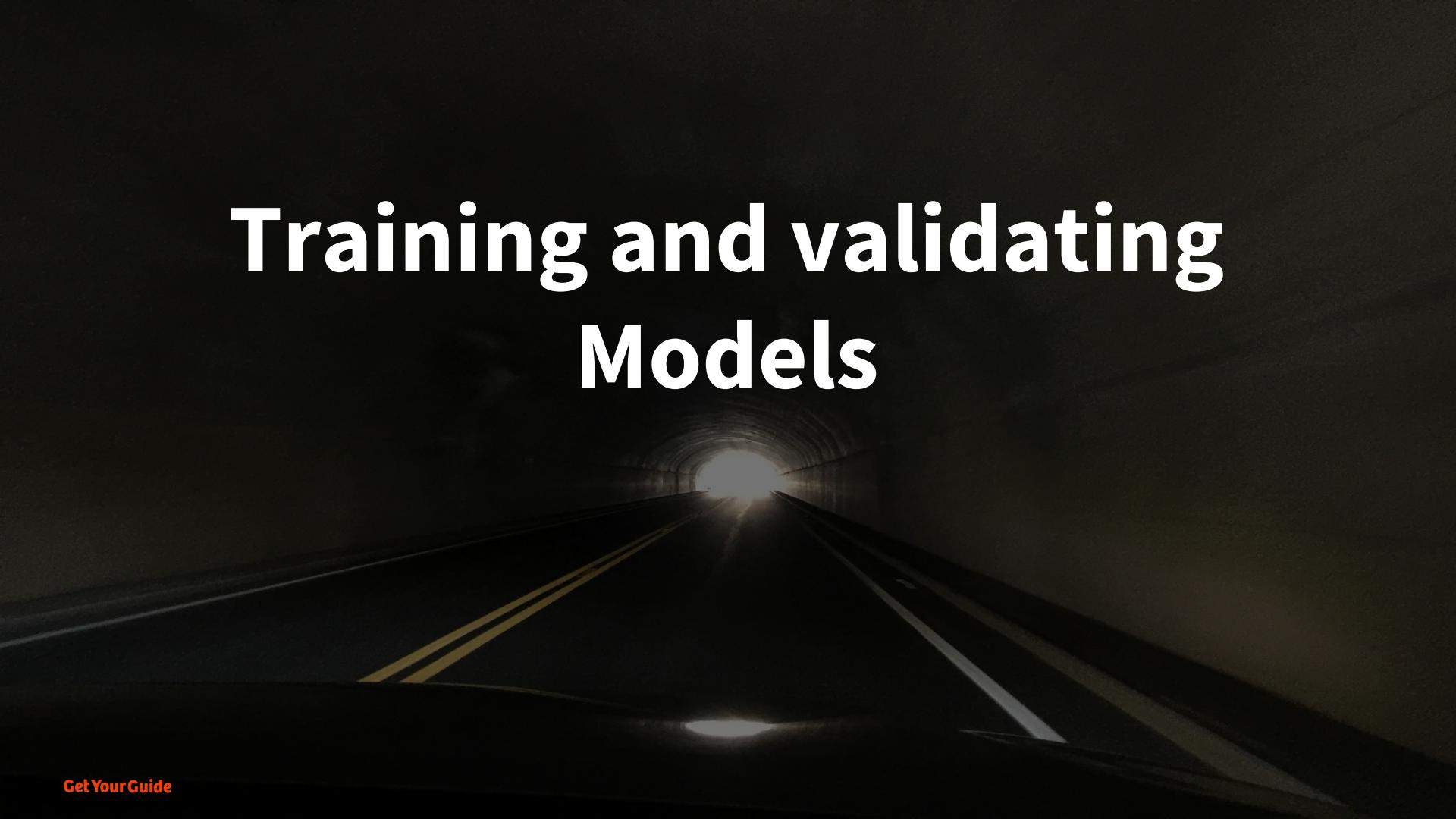
# Our stack

- Elasticsearch
  - LTR Plugin by OpenSource Connections
- RankLib
- Databricks to run our data pipelines
  - Collect features
  - Train and validate models

# New pipeline to collect features



# Training and validating Models



# Goals

- Have a model suitable for location pages
  - relevance + business metrics
- Evaluation metric: **NDCG@10**
- Success (business):  $\uparrow$  **CTR** (Click-Through Rate)
- Constraints
  - Do not include user features

# Best V1 Model

- LambdaMart
- NDCG@10 = 0.9282

Query document	Business metrics	Document
<ul style="list-style-type: none"><li>• Title</li><li>• Highlight</li><li>• Description</li><li>• Best field multi-match</li></ul>	<ul style="list-style-type: none"><li>• Clicks</li><li>• Bookings</li><li>• Impressions</li><li>• CR</li></ul>	<ul style="list-style-type: none"><li>• #Reviews</li><li>• Review rating</li><li>• Deal price</li><li>• Best seller</li></ul>



ONE  
LANE

ROAD SUBJECT TO  
FLOODING  
INDICATORS SHOW DEPTH

ROAD CLOSED

# Best V1 model didn't work



"Eiffel tower skip-the-line ticket"

C U R R E N T R A N K

	<b>Eiffel Tower: Second Floor Priority Access or Summit Ticket</b>	From US\$51.53
	★★★★★ 4918	
	Duration: 2 hours	
	LIKELY TO SELL OUT	
	<b>Eiffel Tower Walk to the 2nd Floor &amp; Optional Summit Access</b>	From US\$40.22
	★★★★★ 130	
	Duration: 1.5 hours	
	LIKELY TO SELL OUT	
	<b>Eiffel Tower Guided Climb with Optional Summit Access</b>	From US\$42.73
	★★★★★ 240	
	Duration: 2 hours	
	LIKELY TO SELL OUT	
	<b>Paris Eiffel Tower Skip-the-Line Tour and Summit Access</b>	From US\$86.72
	★★★★★ 3753	
	Duration: 90 minutes	
	<b>Paris: Eiffel Tower by Night Skip-the-Line Guided Tour</b>	From US\$45.25
	★★★★★ 126	US\$56.56
	Duration: 1.5 hours	
	<b>Skip-the-Line Eiffel Tower Ticket and 2nd Level Tour</b>	From US\$57.81
	★★★★★ 48	
	Duration: 90 minutes	
	<b>Eiffel Tower: Skip-the-Line Summit Tour</b>	From US\$75.28
	★★★★★ 6	
	Duration: 90 minutes	
	<b>Eiffel Tower Priority Access to the Second Floor</b>	From US\$61.58
	★★★★★ 127	
	Duration: 2 hours	
	<b>Skip-the-Line Eiffel Tower Ticket and Small Group Tour</b>	From US\$50.90
	★★★★★ 70	US\$56.56
	Duration: 1.5 hours	

Get Your Guide

M O D E L R A N K

	<b>Skip The Line: 2-Hour Guided Eiffel Tower Tour</b>	From US\$86
	★★★★★ 198	
	Duration: 2 hours	
	LIKELY TO SELL OUT	
	<b>Eiffel Tower Dinner With River Cruise</b>	From US\$209
	★★★★★ 21	
	Duration: 4 hours	
	LIKELY TO SELL OUT	
	<b>Guided Tour to Eiffel Tower 2nd Floor with Priority Access</b>	From US\$55
	No reviews yet	
	Duration: 90 minutes	
	NEW ACTIVITY	
	<b>Paris: City Tour and Skip-the-Line Eiffel Tower Ticket</b>	From US\$273
	Duration: 2 hours	
	<b>Louvre, Notre Dame, Seine Cruise &amp; Eiffel Tower Lunch</b>	From US\$215
	★★★★★ 45	
	Duration: 7.5 hours	
	<b>Eiffel Tower, Hop on Hop Off &amp; Option for Cruise Ticket</b>	From US\$110
	★★★★★ 1	
	Duration: 1 day	
	<b>Paris: Eiffel Tower Dinner &amp; Seine River Cruise</b>	From US\$219
	★★★★★ 20	
	Duration: 4 hours	
	<b>Eiffel Tower: Skip-the-Line Summit Tour</b>	From US\$75.28
	★★★★★ 6	
	Duration: 90 minutes	
	<b>Priority Eiffel Tower Tickets &amp; Audio Guided Paris Bus Tour</b>	From US\$86
	★★★★★ 403	
	Duration: 3 hours	

**We couldn't put in production, shall we give up?**



No, We never give up



# Main lessons learned

- Relevance of results for LP
- Judgement list extraction
- Quality of our queries
- Distribution of judgements

# What is relevance for your business ?

- Our use case: Location pages
  - First point of contact of many visitors
  - Few rank positions to change
  - **Business metrics matter (e.g., revenue)**
- Experts labeling
  - This document is relevant for this query ? **0 - 3**
  - This document is a potential conversion ? 

# Another approach

- Data approach for e-commerces
  - Perceived utility of:
    - search results (Click through rate)
    - product page (Add-to-cart)
  - Overall user satisfaction (Conversion)
  - Business value (Revenue)
- Experts could refine judgements collected from data

Reference: [On Application of Learning to Rank for E-Commerce Search](#) by Santu, Sondhi and Zhai (2017)

# Quality of our queries

- Didn't consider real user query but the keyword search engine matches
- Location part is not relevant for scoring many queries

**"Statue of Liberty boat tour"**

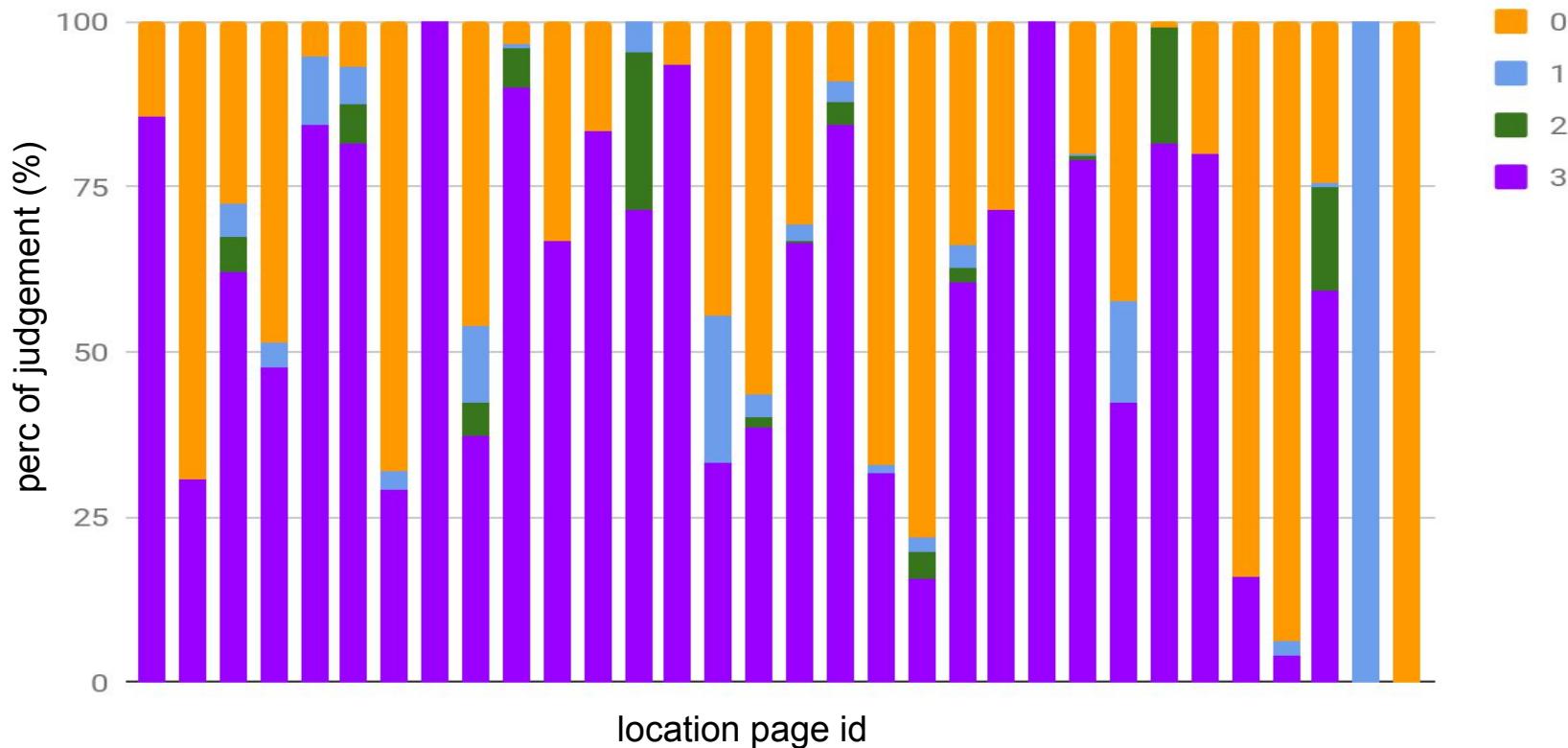


All results  
contain this location

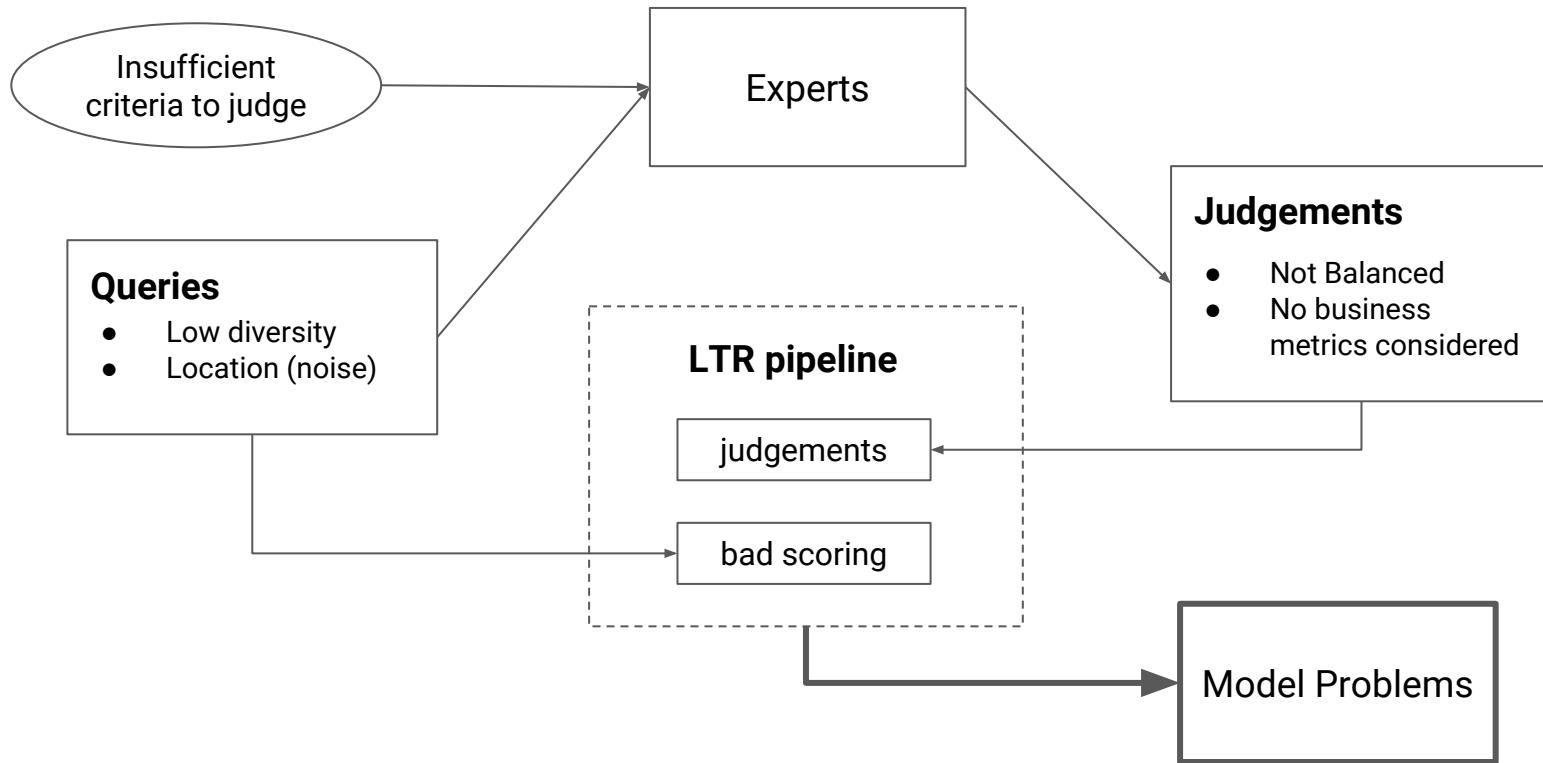


good!

# Distribution of our Judgements per page



# Everything is connected



## Next steps for V2

- Collect judgements from data
- Redefine our criteria for measuring relevance
- Apply LTR in another GYG search features
- Extract the intentions from the keywords
  - Query understanding might help
- Judge the judgements very often

**We hope to turn on V2 and fly**



**Thank you**

# Questions

@AshrafAaref @fmbesson