

# LTR at GetYourGuide Marketplace

A Journey through our experience

Ashraf Aaref and Felipe Besson

June 13th 2018



# 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

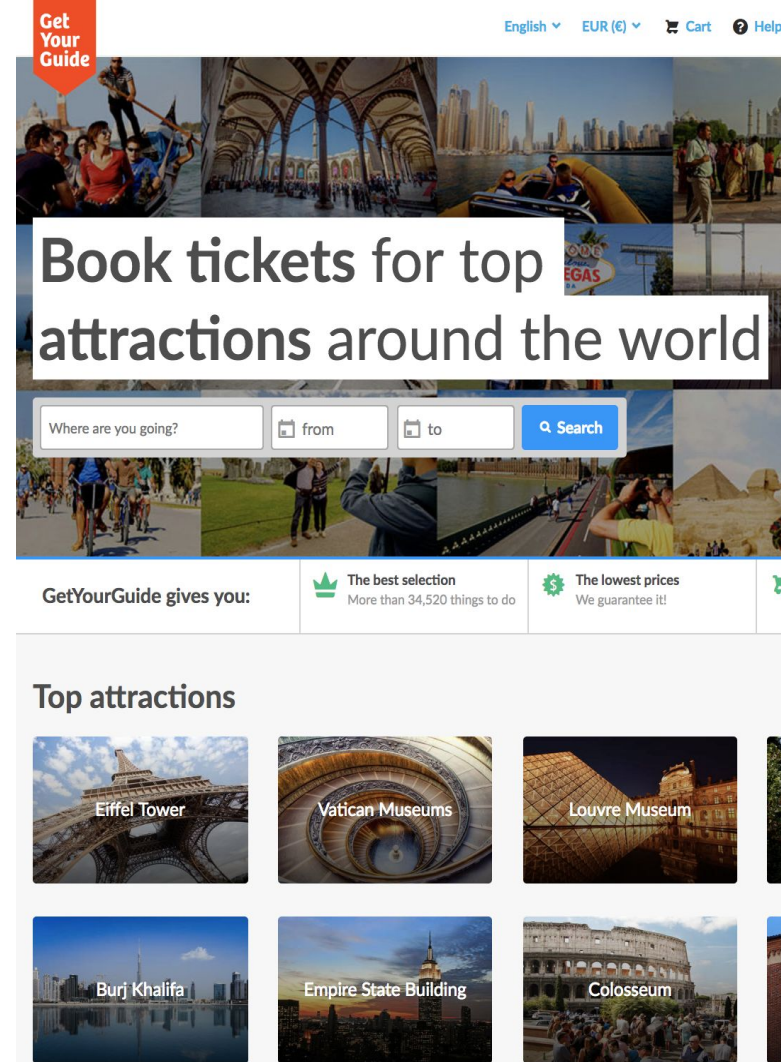
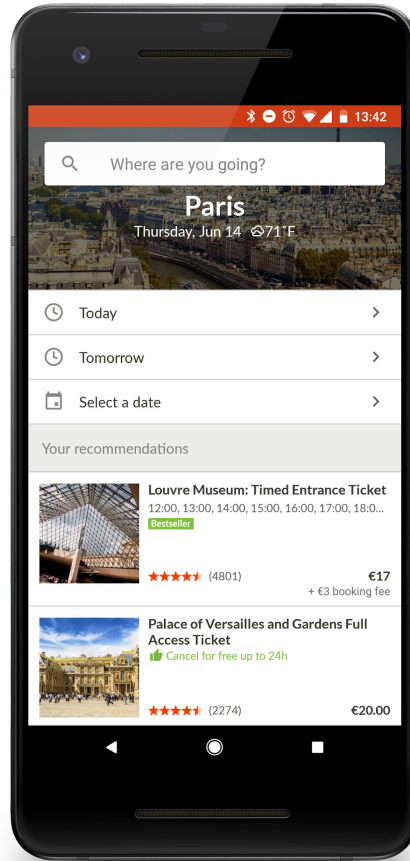
- Location driven
- Discovery

**Rank**

**Business metrics**

+

**Text Relevance**

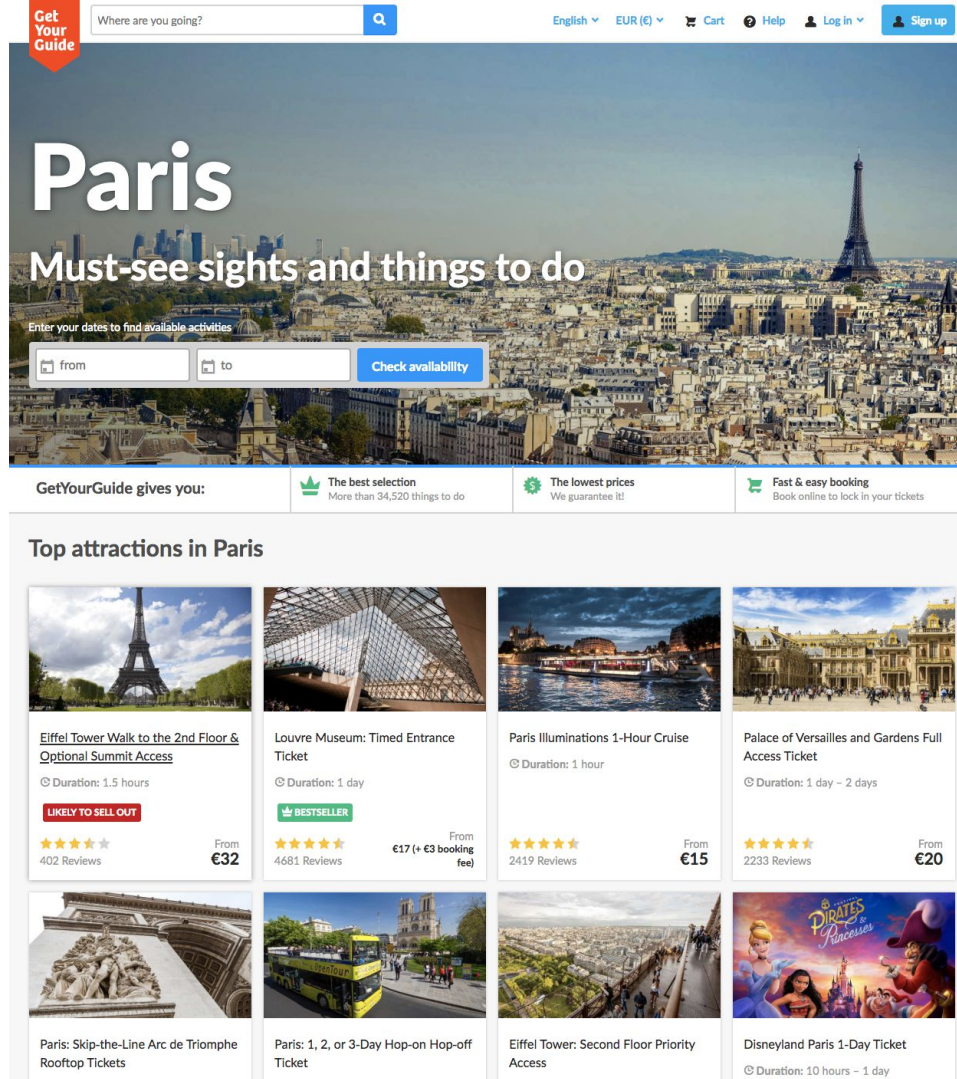


# Location pages (LPs)

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

Rank

Business metrics



The screenshot shows the Get Your Guide website for Paris. At the top, there's a navigation bar with the 'Get Your Guide' logo, a search bar, and links for language (English), currency (EUR), cart, help, and login. The main header features a large image of Paris with the Eiffel Tower, the title 'Paris', and the subtitle 'Must-see sights and things to do'. Below this is a date selection widget with 'from' and 'to' fields and a 'Check availability' button. A banner below the date widget lists three benefits: '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 section 'Top attractions in Paris' displays a grid of eight attractions, each with a thumbnail image, title, duration, reviews, and price. The attractions are: Eiffel Tower Walk to the 2nd Floor & Optional Summit Access (1.5 hours, 402 reviews, €32), Louvre Museum: Timed Entrance Ticket (1 day, 4681 reviews, €17), Paris Illuminations 1-Hour Cruise (1 hour, 2419 reviews, €15), Palace of Versailles and Gardens Full Access Ticket (1 day - 2 days, 2233 reviews, €20), 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 (10 hours - 1 day).

Get Your Guide

Where are you going?

English ▼ EUR (€) ▼ Cart Help Log in Sign up

# Paris









## Must-see sights and things to do

Enter your dates to find available activities

GetYourGuide gives you:

- The best selection**  
More than 34,520 things to do
- The lowest prices**  
We guarantee it!
- Fast & easy booking**  
Book online to lock in your tickets

### Top attractions in Paris

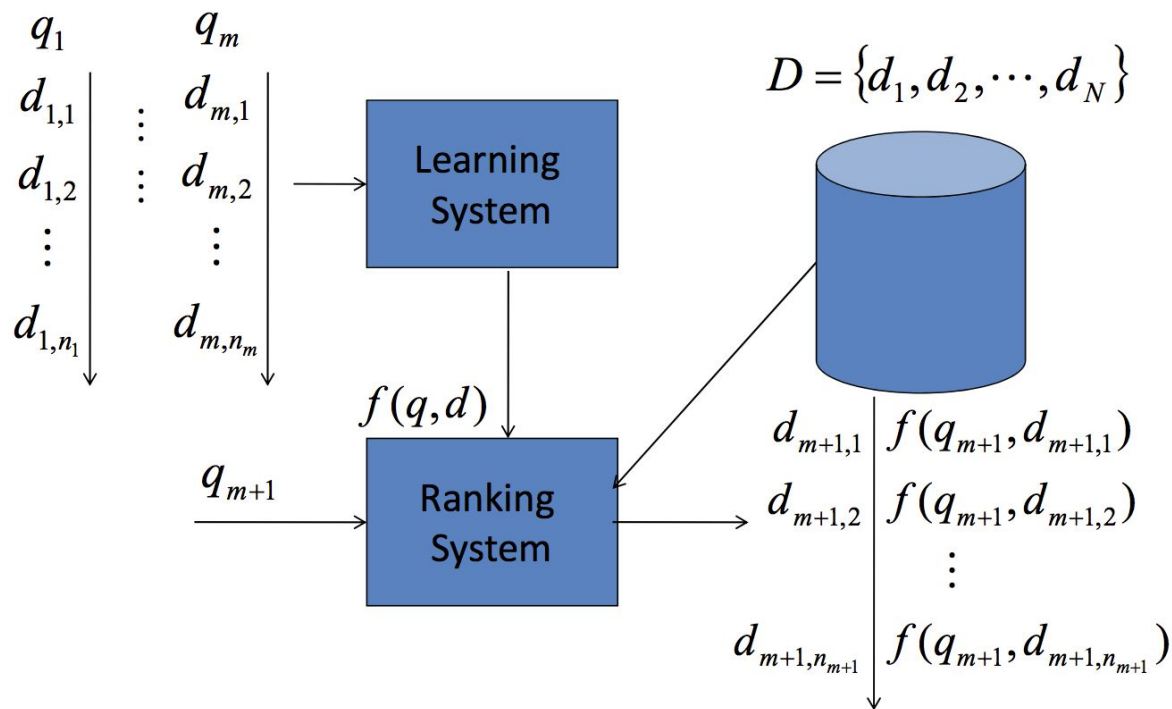
 <b>Eiffel Tower Walk to the 2nd Floor &amp; Optional Summit Access</b> ⌚ Duration: 1.5 hours <b>LIKELY TO SELL OUT</b> ★★★★☆ 402 Reviews From <b>€32</b>	 <b>Louvre Museum: Timed Entrance Ticket</b> ⌚ Duration: 1 day <b>BESTSELLER</b> ★★★★★ 4681 Reviews From <b>€17 (+ €3 booking fee)</b>	 <b>Paris Illuminations 1-Hour Cruise</b> ⌚ Duration: 1 hour ★★★★★ 2419 Reviews From <b>€15</b>	 <b>Palace of Versailles and Gardens Full Access Ticket</b> ⌚ Duration: 1 day - 2 days ★★★★★ 2233 Reviews From <b>€20</b>
 <b>Paris: Skip-the-Line Arc de Triomphe Rooftop Tickets</b>	 <b>Paris: 1, 2, or 3-Day Hop-on Hop-off Ticket</b>	 <b>Eiffel Tower: Second Floor Priority Access</b>	 <b>Disneyland Paris 1-Day Ticket</b> ⌚ Duration: 10 hours - 1 day

# 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"**



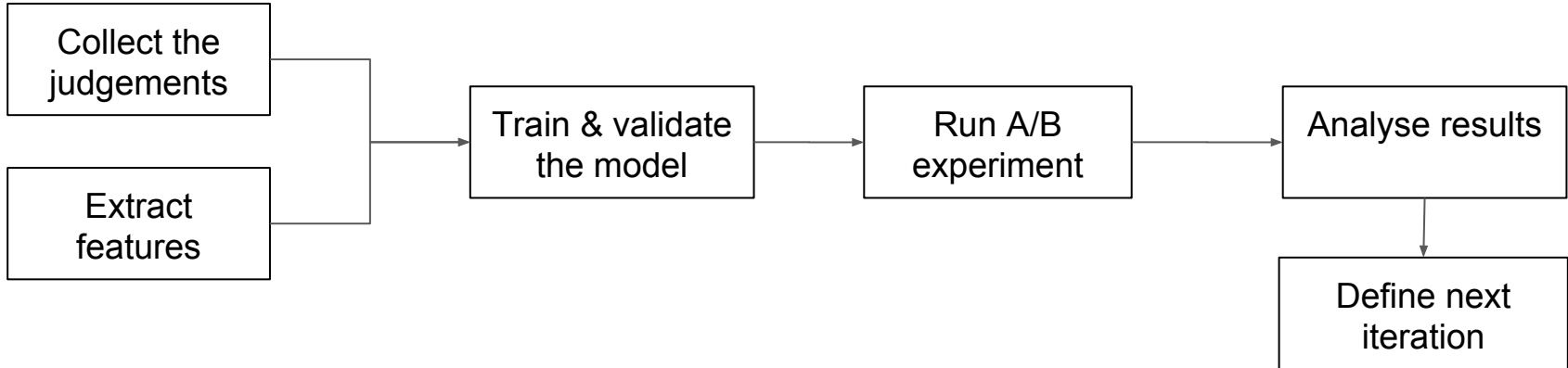
**location**



**intention**

# MVP mindset

Follow the standard steps of a LTR solution



**We started the journey!**








# Judgement List

q = "Eiffel Tower restaurant"

## Document

## Judgement

	<p><b>Dinner at 58 Tour Eiffel with Priority Lift Tickets</b></p> <p>★★★★★ 701 Reviews</p> <p>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...</p> <p>⌚ Duration: 2 hours – 2.5 hours</p> <p>From <b>US\$114.14</b></p>
	<p><b>Eiffel Lunch, 2nd Floor Ticket and Seine River Cruise</b></p> <p>★★★★★ 16 Reviews</p> <p>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...</p> <p>⌚ Duration: 4 hours</p> <p>From <b>US\$133.56</b></p>
	<p><b>Eiffel Tower: Second Floor Priority Access</b></p> <p>★★★★★ 5484 Reviews <b>LIKELY TO SELL OUT</b></p> <p>Save precious time at the Eiffel Tower with a priority-access ticket to the second level.</p> <p>⌚ Duration: 2 hours</p> <p>From <b>US\$49.79</b></p>
	<p><b>FlyView: Fly over Paris in Virtual Reality</b></p> <p>★★★★★ 6 Reviews</p> <p>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.</p> <p>⌚ Duration: 31 minutes</p> <p><del>US\$18.21</del> From <b>US\$14.57</b></p>
	<p><b>Tower of London Ticket with Crown Jewels Exhibition</b></p> <p>★★★★★ 1495 Reviews <b>LIKELY TO SELL OUT</b></p> <p>Visit the Tower of London and see the famous building that has served as a fortress, palace, and prison.</p> <p>⌚ Duration: 60 minutes – 6 hours</p> <p>From <b>US\$37.04</b></p>

3

3

2

1

0

# 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
  - $NDCG@7 = \mathbf{0.55}$



# 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>



A perspective view of a road entering a dark tunnel. The road has double yellow lines and leads towards a bright light at the tunnel's exit. The walls of the tunnel are rocky and uneven. The text "How to collect these features ?" is overlaid in the center in a large, white, bold font.

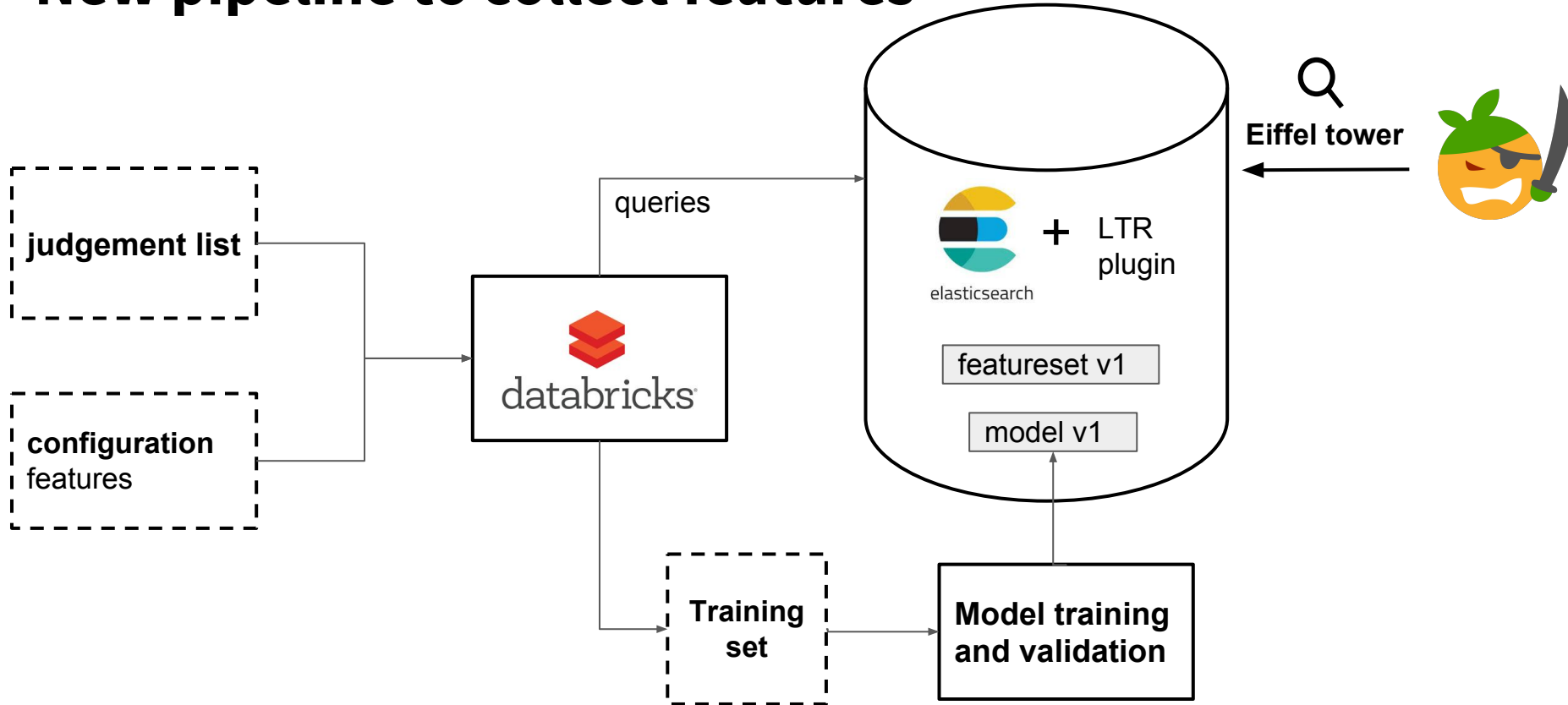
# How to collect these features ?

# 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): **↑ CTR** (Click-Through Rate)
- Constraints
  - Do not include user features

# Best V1 Model

- LambdaMart
- $\text{NDCG@10} = 0.9282$

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>








# Best V1 model didn't work




"Eiffel tower skip-the-line ticket"

C  
U  
R  
R  
E  
N  
T  
  
R  
A  
N  
K




**Eiffel Tower: Second Floor Priority Access or Summit Ticket**  
★★★★★ 4918  
Duration: 2 hours  
**LIKELY TO SELL OUT**

From  
**US\$51.53**




**Eiffel Tower Walk to the 2nd Floor & Optional Summit Access**  
★★★★★ 130  
Duration: 1.5 hours  
**LIKELY TO SELL OUT**

From  
**US\$40.22**



**Eiffel Tower Guided Climb with Optional Summit Access**  
★★★★★ 240  
Duration: 2 hours  
**LIKELY TO SELL OUT**

From  
**US\$42.73**




**Paris Eiffel Tower Skip-the-Line Tour and Summit Access**  
★★★★★ 3753  
Duration: 90 minutes

From  
**US\$86.72**



**Paris: Eiffel Tower by Night Skip-the-Line Guided Tour**  
★★★★★ 126  
Duration: 1.5 hours

From  
**US\$45.25**  
~~US\$56.56~~




**Skip-the-Line Eiffel Tower Ticket and 2nd Level Tour**  
★★★★★ 48  
Duration: 90 minutes

From  
**US\$57.81**




**Eiffel Tower: Skip-the-Line Summit Tour**  
★★★★★ 6  
Duration: 90 minutes

From  
**US\$75.28**



**Eiffel Tower Priority Access to the Second Floor**  
★★★★★ 127  
Duration: 2 hours

From  
**US\$61.58**




**Skip-the-Line Eiffel Tower Ticket and Small Group Tour**  
★★★★★ 70  
Duration: 1.5 hours

From  
**US\$50.90**  
~~US\$56.56~~




**Skip The Line: 2-Hour Guided Eiffel Tower Tour**  
★★★★★ 198  
Duration: 2 hours  
**LIKELY TO SELL OUT**

From  
**US\$86**




**Eiffel Tower Dinner With River Cruise**  
★★★★★ 21  
Duration: 4 hours  
**LIKELY TO SELL OUT**

From  
**US\$209**




**Guided Tour to Eiffel Tower 2nd Floor with Priority Access**  
No reviews yet  
Duration: 90 minutes  
**NEW ACTIVITY**

Intro Special  
From  
**US\$55**




**Paris: City Tour and Skip-the-Line Eiffel Tower Ticket**  
Duration: 2 hours

From  
**US\$273**



**Louvre, Notre Dame, Seine Cruise & Eiffel Tower Lunch**  
★★★★★ 45  
Duration: 7.5 hours

From  
**US\$215**




**Eiffel Tower, Hop on Hop Off & Option for Cruise Ticket**  
★★★★★ 1  
Duration: 1 day

From  
**US\$110**




**Paris: Eiffel Tower Dinner & Seine River Cruise**  
★★★★★ 20  
Duration: 4 hours

From  
**US\$219**



**Eiffel Tower: Skip-the-Line Summit Tour**  
★★★★★ 6  
Duration: 90 minutes

From  
**US\$75.28**



**Priority Eiffel Tower Tickets & Audio Guided Paris Bus Tour**  
★★★★★ 403  
Duration: 3 hours

From  
**US\$86**

M  
O  
D  
E  
L  
  
R  
A  
N  
K

**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"



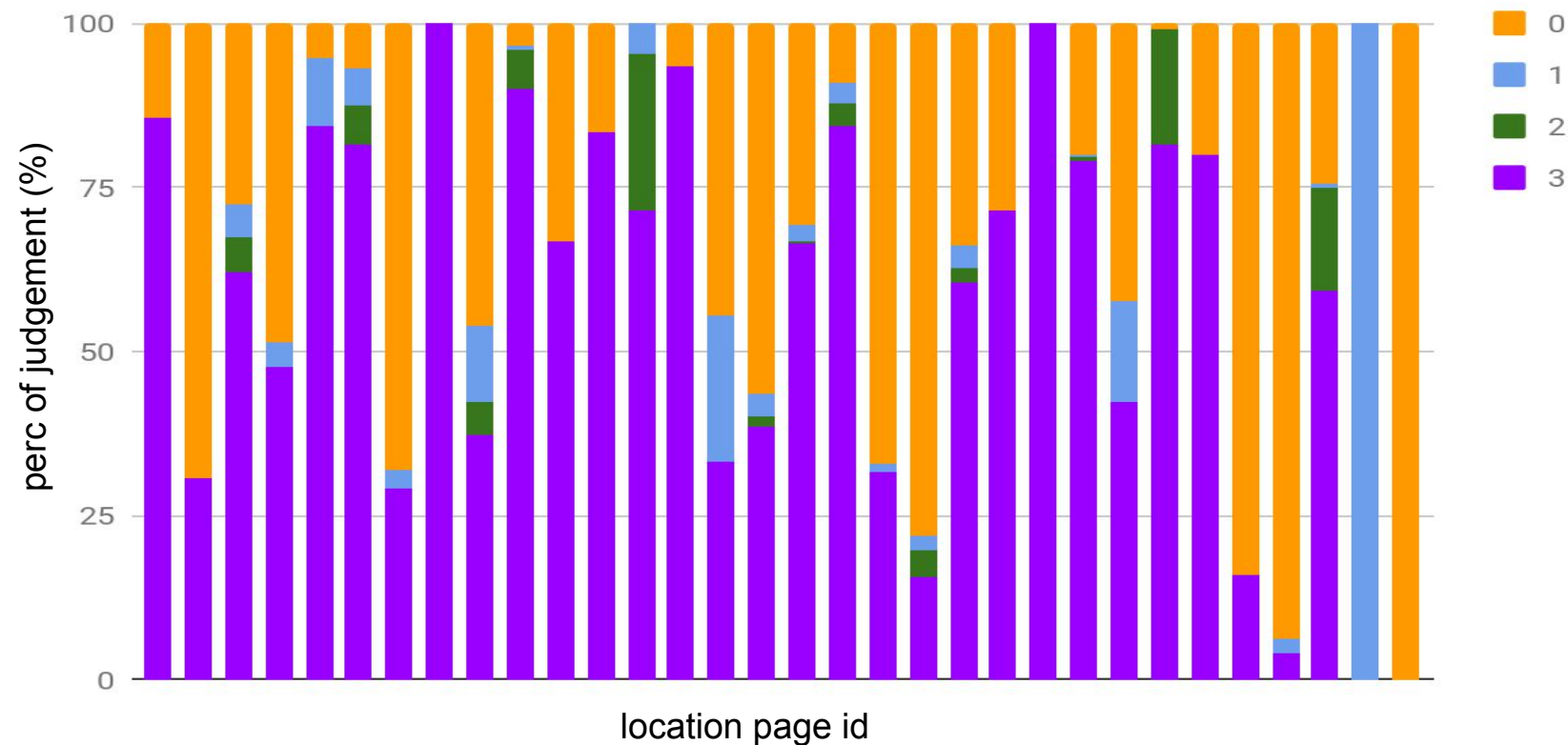
All results  
contain this location



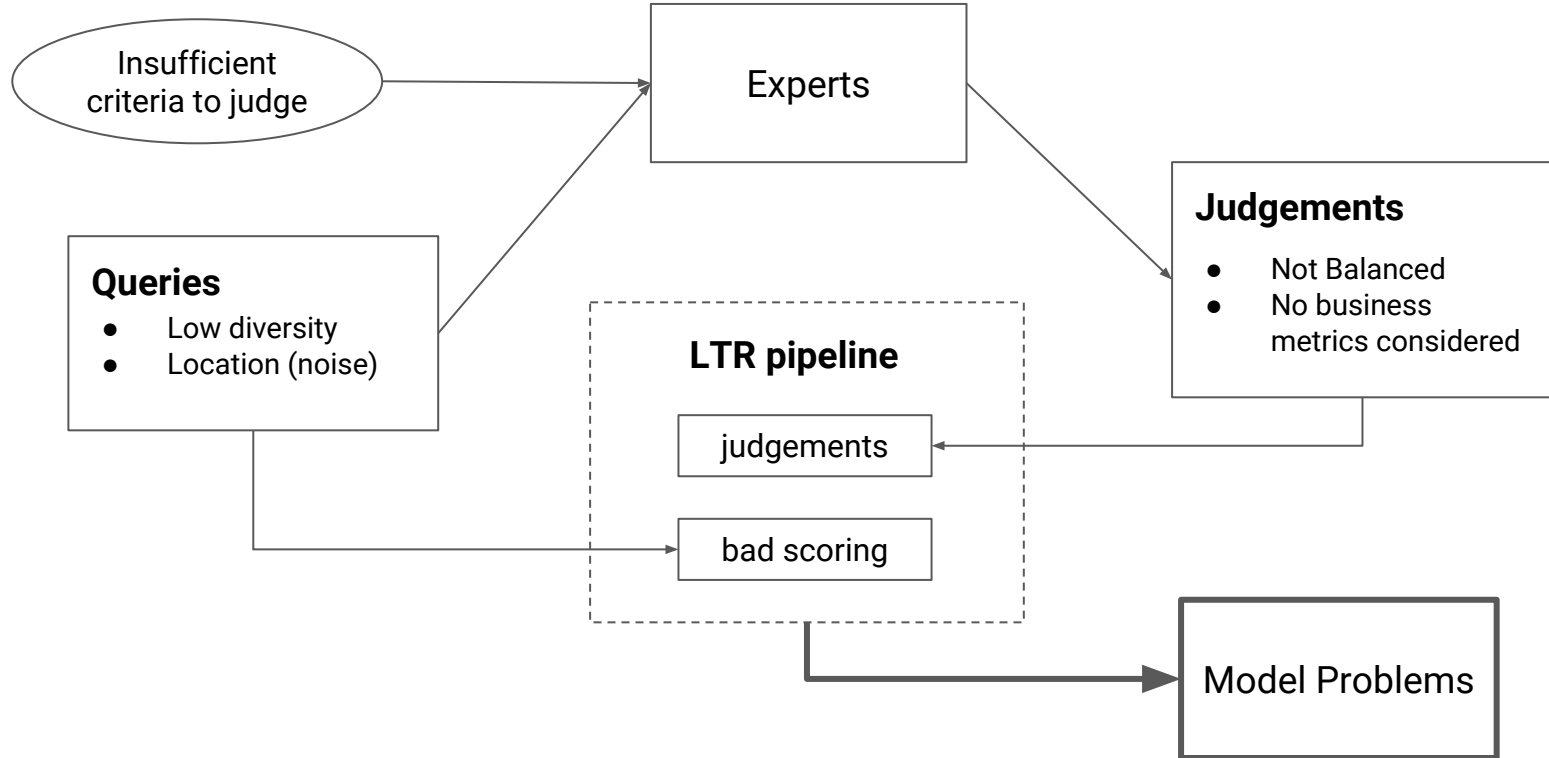
good!



# 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