Breaking the Ice: Solving the Item Cold Start Problem in E-commerce Search

Evgeniia Trufanova

MICES EU 2023

Delivery Hero

Hello! 👋

DeliveryTech

Data Scientist with 4 years of experience:

- **Startup**
  - Fashion AI
- **University thesis on ML in medicine**
  - DSS for diabetes management
- **Banking industry**
  - NLP
  - AutoML tools
- **Currently at Delivery Hero**
  - Search Ranking

Who we are

Delivery Hero

- Food ordering & delivery
- Restaurants & Quick commerce
- 70+ countries

2.8 billion orders
*2021 public data

500 000 restaurants
*2020 public data

Brands:

foodpanda

Glovo

PedidosYa

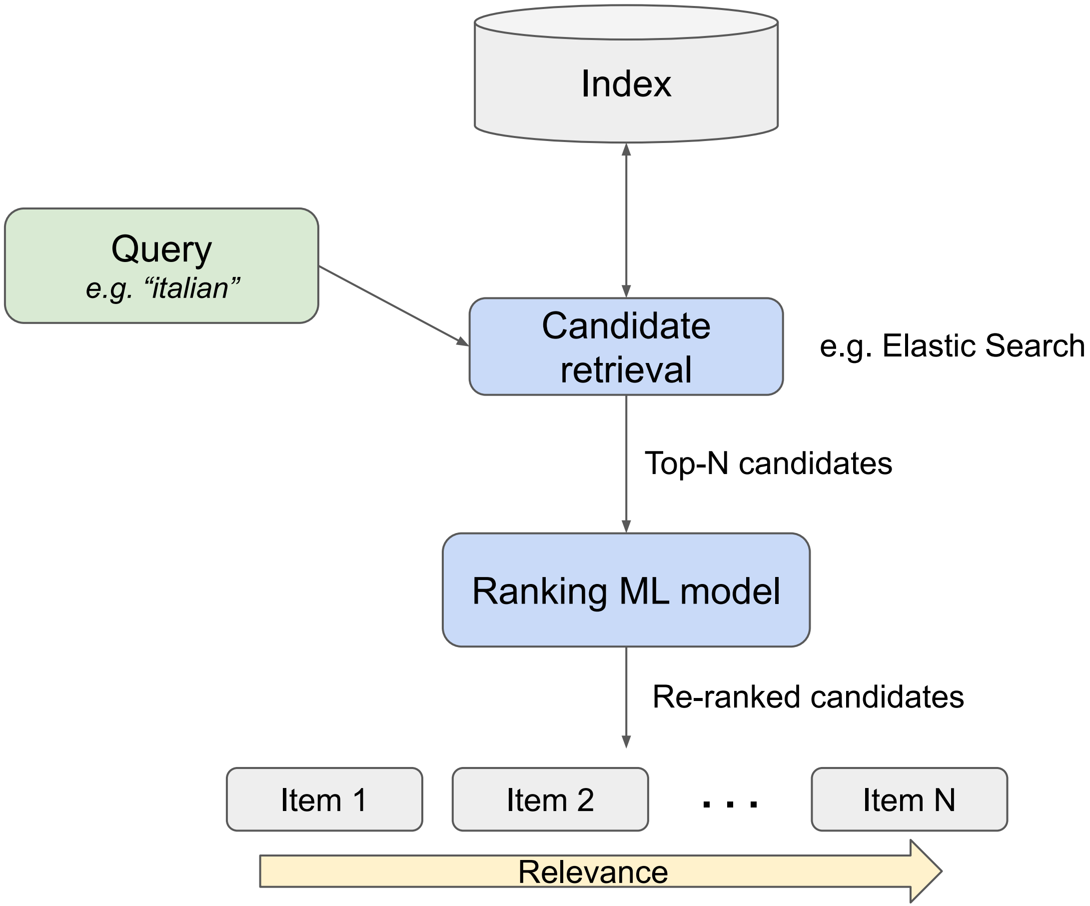
Yemeksepeti

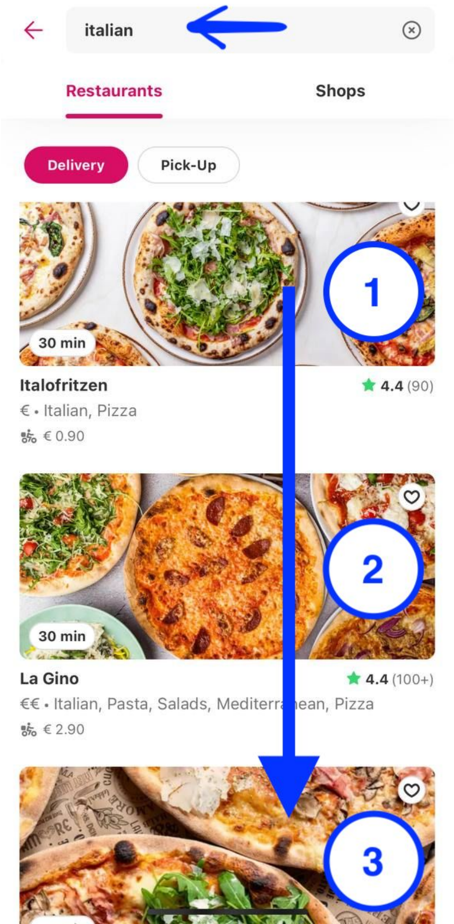
foodora

talabat

E-commerce Search Ranking

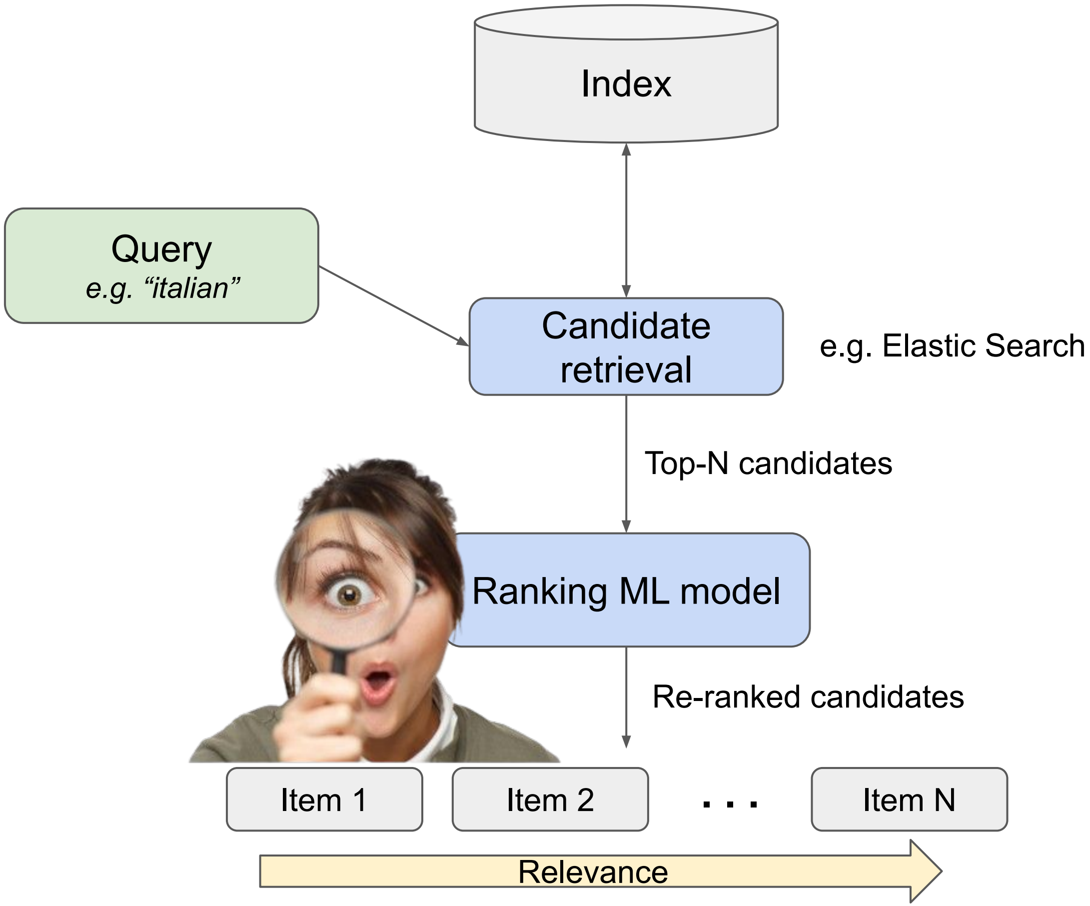
DeliveryTech

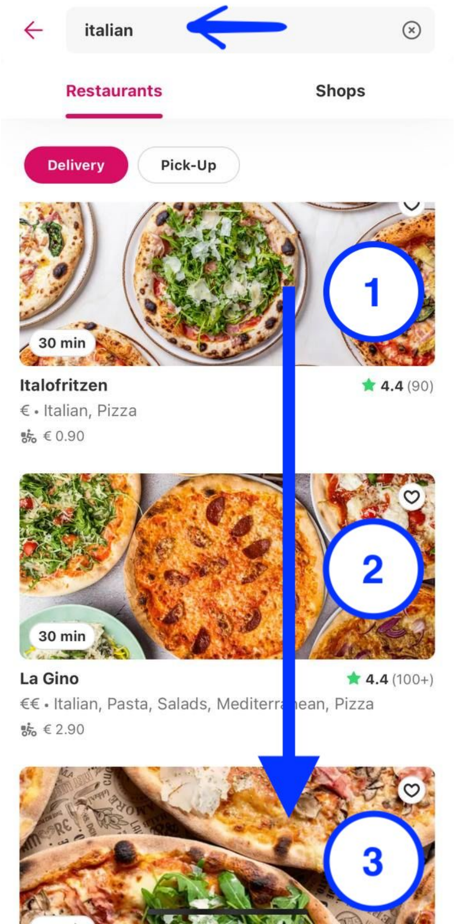




E-commerce Search Ranking

DeliveryTech





Ranking ML model

Features:

Importance

Behavioral features, past interactions
Popularity, rating, CTR, number of past orders & clicks

Textual features
String or embedding similarity

Context features
User context, time context, personalization features

Price features
Item price, delivery fee

Multimodal features
Image-based similarity

Ranking ML model

Features:

Importance

Behavioral features, past interactions
Popularity, rating, CTR, number of past orders & clicks

What about new restaurants?

Textual features
String or embedding similarity

Context features
User context, time context, personalization features

Price features
Item price, delivery fee

Multimodal features
Image-based similarity

New restaurant challenges

New restaurant signing up on Delivery Hero platform:

New restaurant challenges

New restaurant signing up on Delivery Hero platform:

0 past orders

New restaurant challenges

New restaurant signing up on Delivery Hero platform:

0 past orders

No trackable CVR

0 past clicks

New restaurant challenges

New restaurant signing up on Delivery Hero platform:

0 past orders

No trackable CVR

0 past clicks

Position in search results: 47

New restaurant challenges

New restaurant signing up on Delivery Hero platform:

0 past orders

No trackable CVR

0 past clicks

Position in search results: 47

NO orders
coming from our platform

New restaurant challenges

New restaurant signing up on Delivery Hero platform:

0 past orders

No trackable CVR

Position in search results: 47

NO orders
coming from our platform

0 past clicks

New restaurant challenges

New restaurant signing up on Delivery Hero platform:

0 past orders → No trackable CVR → Position in search results: 47 → NO orders coming from our platform → 0 past clicks → 0 past orders

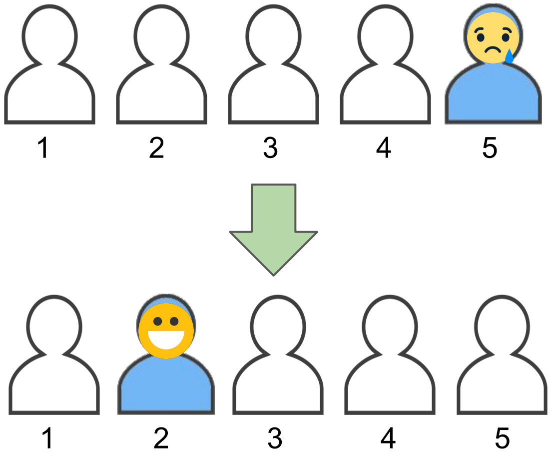
NO orders
coming from our platform

Poor experience for new restaurants = NO growth

Cold start problem 🥶

Context:

- Ranking ML model relies on past interactions
- New items do not have any past interactions
- New items are displayed on lower positions and receive few clicks and orders

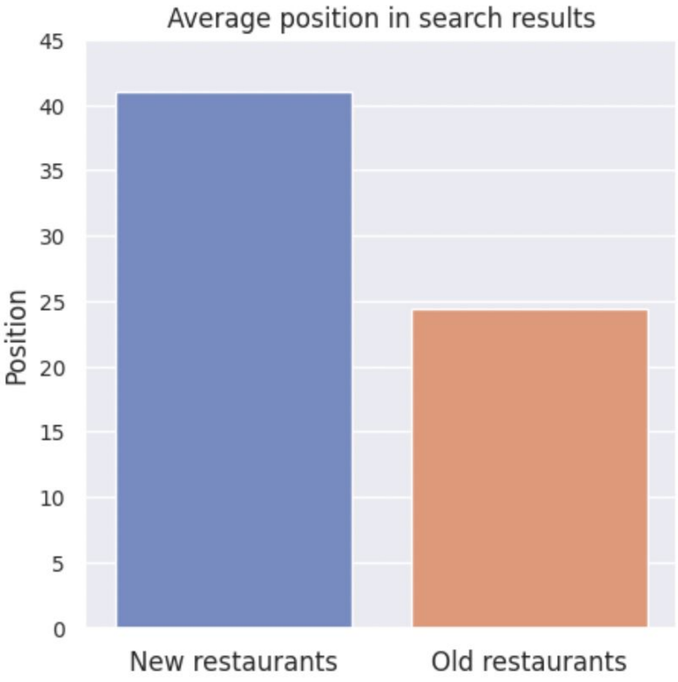


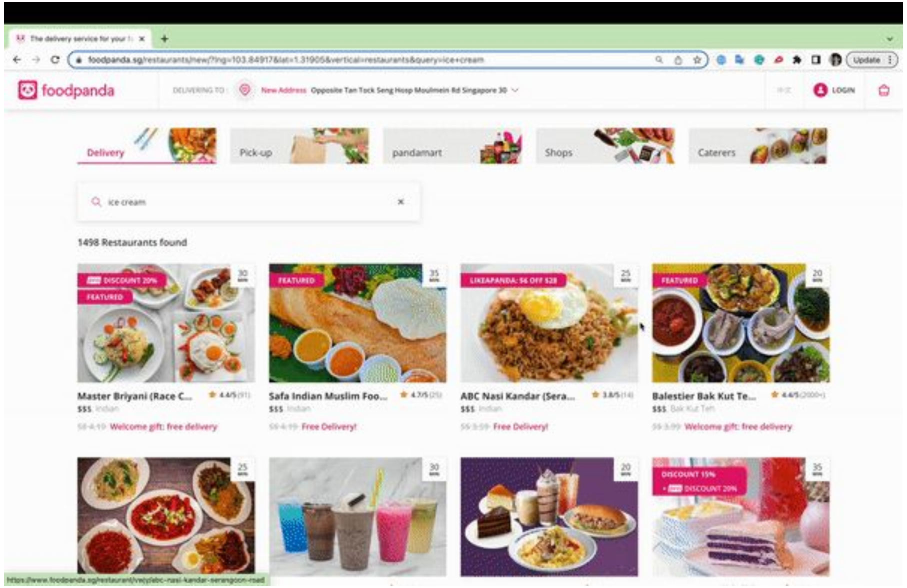
Goal:

- Increase visibility & traffic for new items without quality loss for others

Cold start problem 🥶

Cold start problem at Delivery Hero:



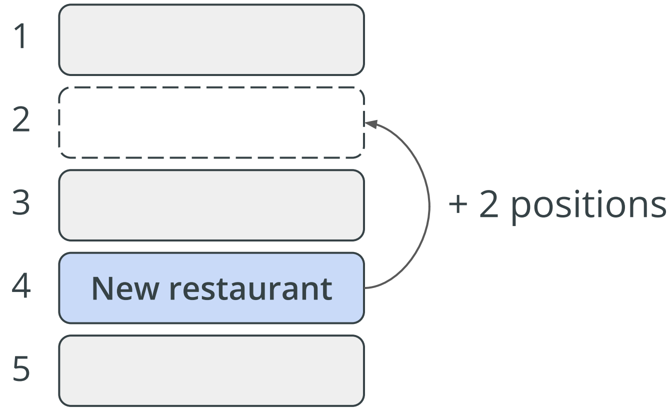


Approach 1: Artificial Boosting

DeliveryTech

Boost cold start restaurants

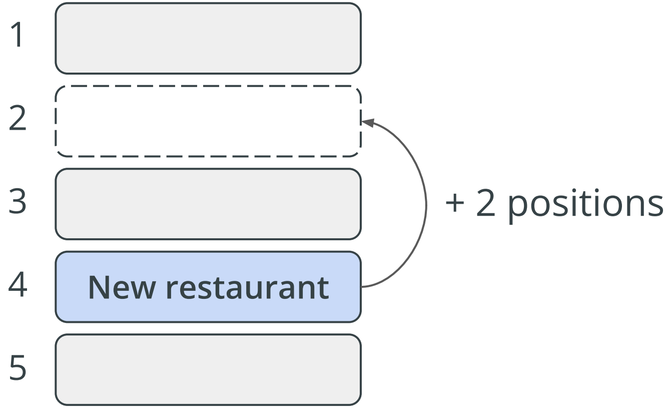
Artificially move new restaurants N positions higher or use a multiplier for their score



Approach 1: Artificial Boosting

Boost cold start restaurants

Artificially move new restaurants N positions higher or use a multiplier for their score



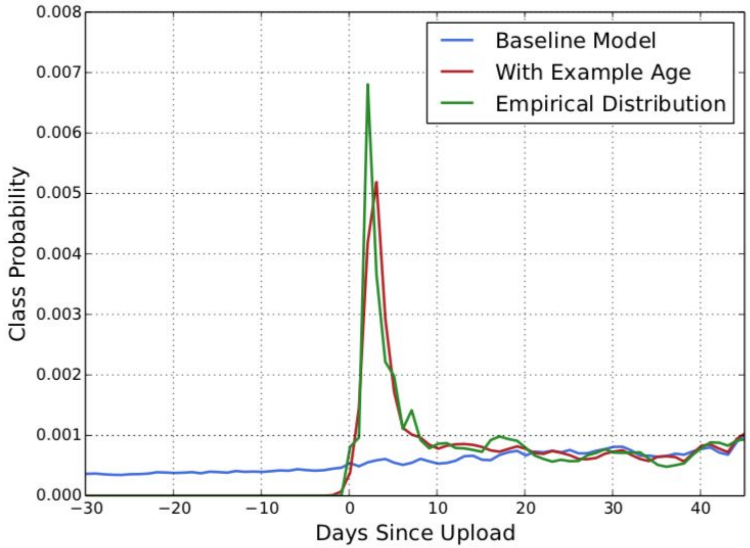
Result:

Neutral A/B test

Approach 2: Cold Start Features

Use cold start as a feature

Use “restaurant age” / “item age” / “days since upload” as a feature for ML model

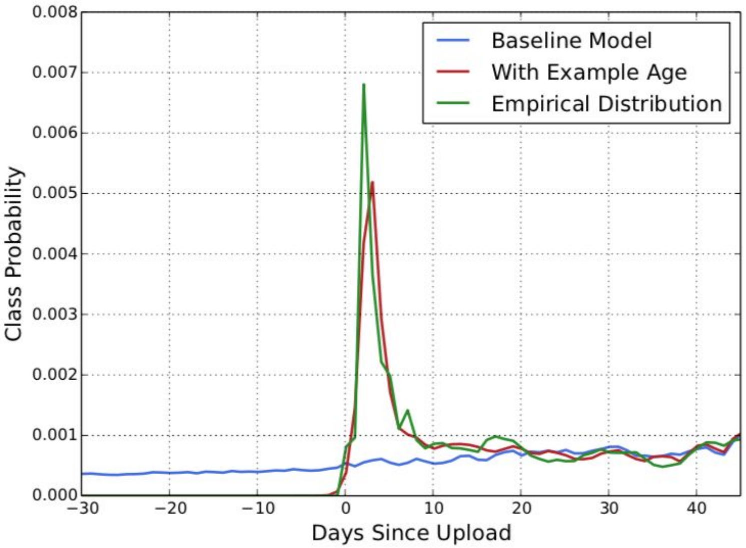


Deep Neural Networks for YouTube Recommendations
https://research.google/pubs/pub45530/

Approach 2: Cold Start Features

Use cold start as a feature

Use “restaurant age” / “item age” / “days since upload” as a feature for ML model



Deep Neural Networks for YouTube Recommendations
https://research.google/pubs/pub45530/

Result: 

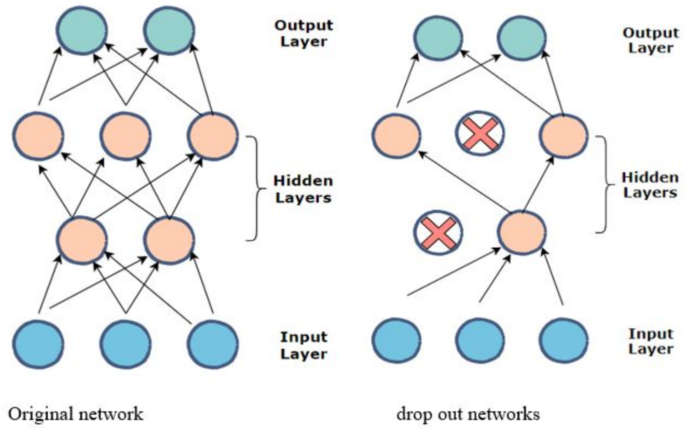
No improvement in offline metrics

Approach 3: DropoutNet

DropoutNet
Inspired by neural networks dropout

During the training, randomly apply *input dropout* to past interaction features: set them to 0 for some fraction of restaurants

This way model does not rely on past interactions too much and learns to generalize to cold start while preserving warm start accuracy



DropoutNet: Addressing Cold Start in Recommender Systems
https://www.cs.toronto.edu/~mvolkovs/nips2017_deepcf.pdf

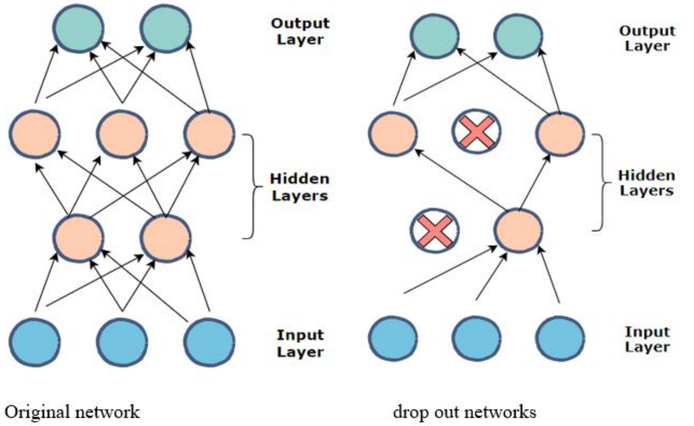
Approach 3: DropoutNet

DropoutNet

Inspired by neural networks dropout

During the training, randomly apply *input dropout* to past interaction features: set them to 0 for some fraction of restaurants

This way model does not rely on past interactions too much and learns to generalize to cold start while preserving warm start accuracy



Result: 

No improvement in offline metrics again

DropoutNet: Addressing Cold Start in Recommender Systems

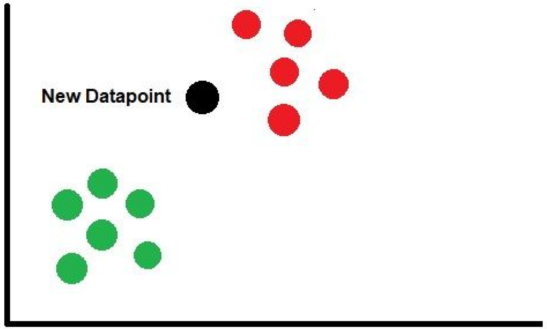
https://www.cs.toronto.edu/~mvolkovs/nips2017_deepcf.pdf

Approach 4: Fill Missing Interactions

*Delivery*Tech

Fill missing past interaction features

- **Representative based**
  - Assumption: new restaurant's interaction counts would be close to similar "old" restaurants.
  - Thus we can approximate its interaction feature values with average of those of similar restaurants.



- **Predicting priors**

  Create a separate model to predict interaction features for new cold start restaurants.

  (This model should be trained on "old" restaurants)

Treating Cold Start in Product Search by Priors
https://assets.amazon.science/55/9d/68fe373d4379878073d66611141b/treating-cold-start-in-product-search-by-priors.pdf

Approach 4: Fill Missing Interactions

Steps to predict interaction features:

1. Create a separate model to predict order & click counts based on some query and item features
2. Use this model to fill missing interaction features for cold start items
3. Apply the normal ranking ML model

	click_count	order_count
Item A	10	2
Item B	235	53
Item C	**–**	**–**
Item D	24	10

Click&order count predictor

	click_count	order_count
Item A	10	2
Item B	235	53
Item C	**15**	**3**
Item D	24	10

Ranking ML model

Approach 5: Hybrid Ranking

DeliveryTech

Hybrid ranking

- Have a separate ranking algorithm that would not not rely on past interaction features (e.g. model trained on content features only)
- Fall back for that algorithm for cold start restaurants

Approach 5: Hybrid Ranking

Elastic Search rank “hack”

1. Keep ES rank for cold start restaurants
2. The rest of the restaurants are ranked and ordered with the normal ranking model

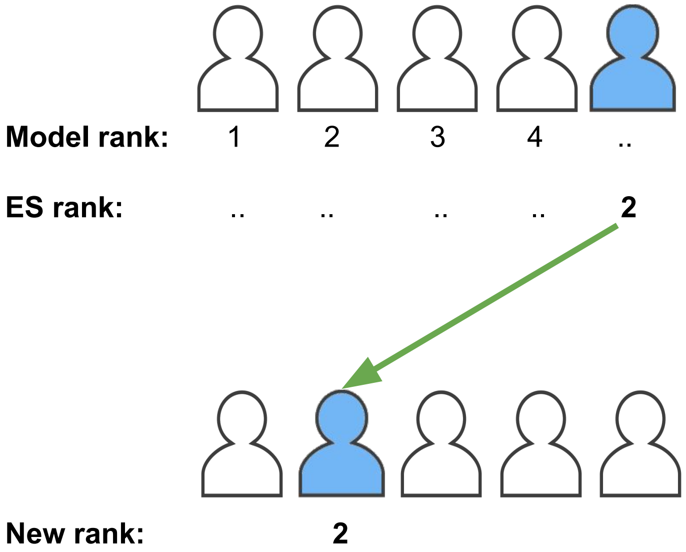
Model rank:	1	2	3	4	..
ES rank:	..	..	..	..	**2**

Approach 5: Hybrid Ranking

DeliveryTech

Elastic Search rank "hack"

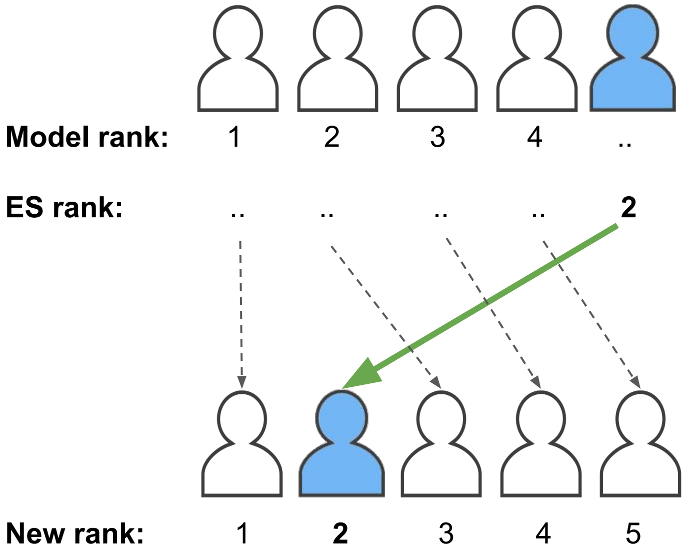
1. Keep ES rank for cold start restaurants
2. The rest of the restaurants are ranked and ordered with the normal ranking model



Approach 5: Hybrid Ranking

Elastic Search rank "hack"

1. Keep ES rank for cold start restaurants
2. The rest of the restaurants are ranked and ordered with the normal ranking model

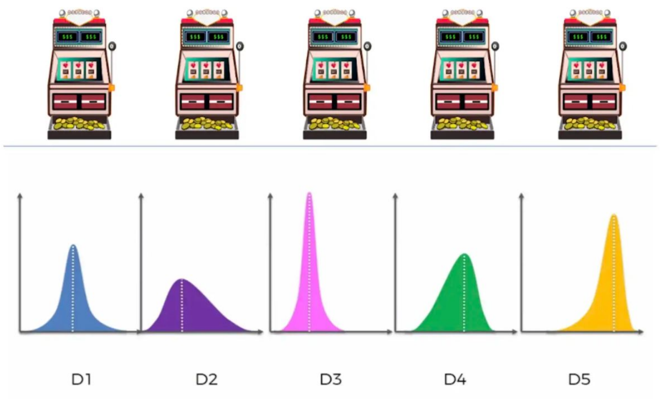


Approach 5a: Multi-Armed Bandits

DeliveryTech

Multi-Armed Bandits

- Balance between exploration and exploitation
- Preferring old vendors - exploitation
- Preferring new vendors - exploration



- **Scenario:** Pull machine k → sample from ***unknown*** reward distribution D_k → observe reward.
- **Problem:** Given a finite number of pulls T, how can I optimize my winnings?
- How much should I **explore**? How much should I **exploit**?

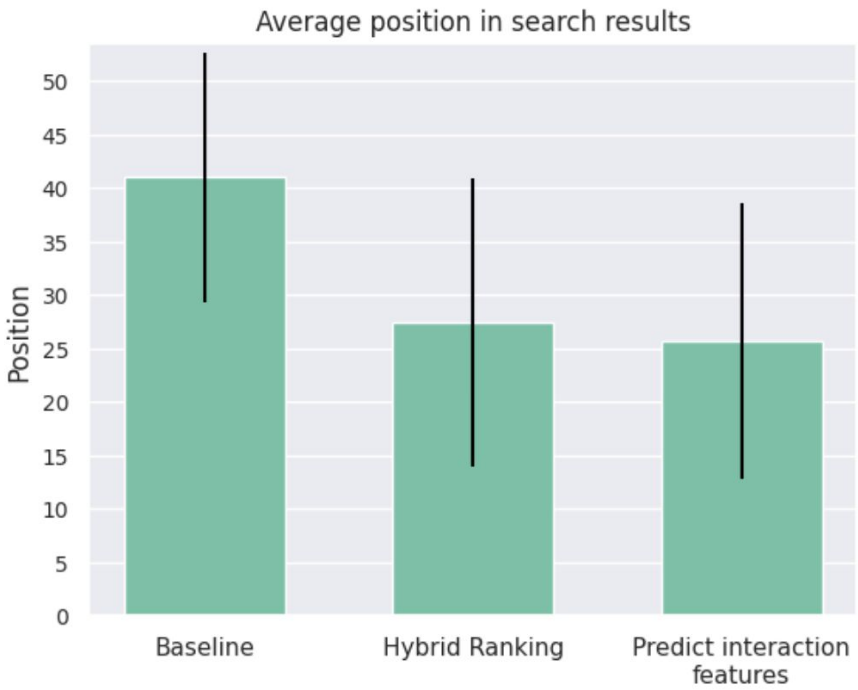
Our Results

Overall nDCG
- Non-significant difference (<0.1%)

Cold start sessions nDCG
- Hybrid Ranking: **+39.9%**
- Predict interaction features: **+52.1%**

Average position of new restaurants
- Hybrid Ranking: **27**
- Predict interaction features: **26**
- Baseline: **41**

A/B test is ongoing



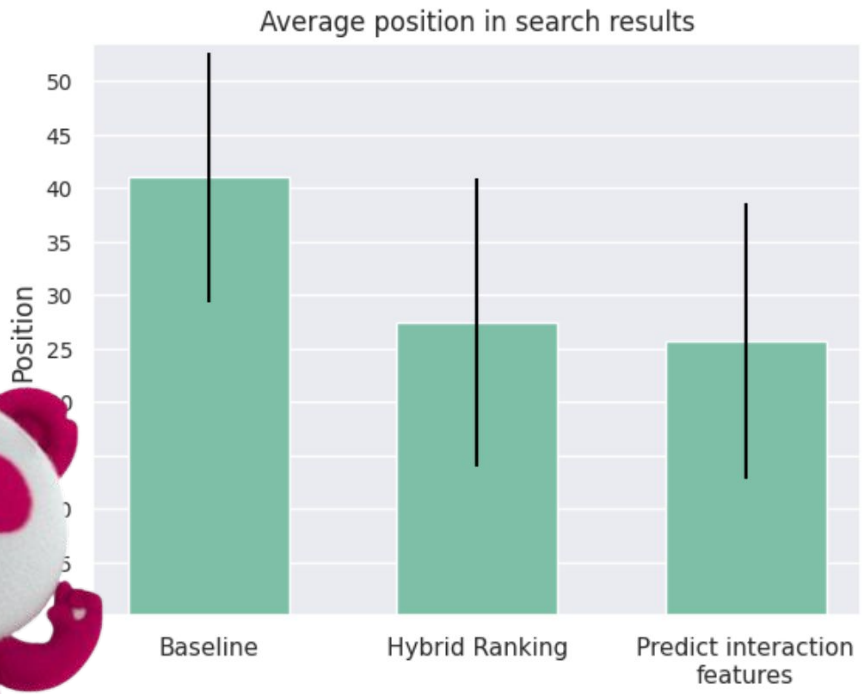
Our Results

Overall nDCG
- Non-significant difference (<0.1%)

Cold start sessions nDCG
- Hybrid Ranking: **+39.9%**
- Predict interaction features: **+52.1%**

Average position of new restaurants
- Hybrid Ranking: **27**
- Predict interaction features: **26**
- Baseline: **41**

A/B test is ongoing



Thank you!

Delivery Hero Tech Blog: tech.deliveryhero.com

Me: linkedin.com/in/evgeniia-trufanova