Building Robust and Scalable Recommendation Systems for Online Food Delivery

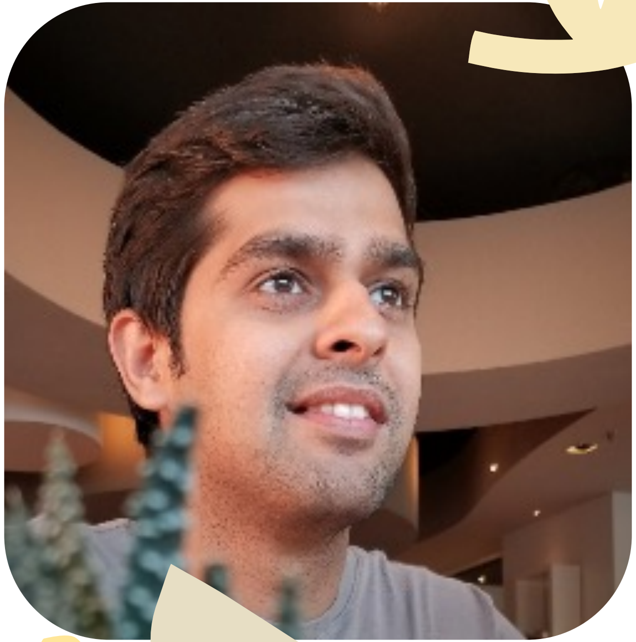
Raghav Bali | Vishal Natani

Agenda

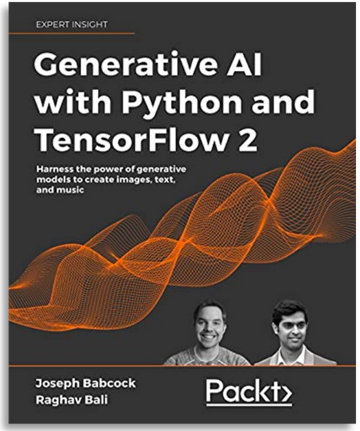
- Recommendation Systems
- Overview of DS Workflow
  - EDA
  - Models
  - Evaluation Metrics:
  - A/B Testing
- Architecture
  - Reference System
- DHRD

Raghav Bali

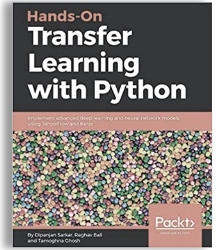
- Staff Data Scientist @ Delivery Hero
- Over a decade-long journey of pioneering enterprise-level solutions. Harnessing Machine Learning, Deep Learning, Computer Vision, NLP and Recommendation Systems

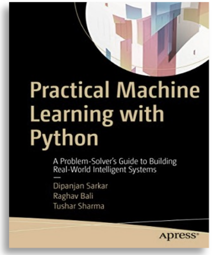


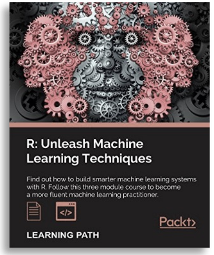
Publications

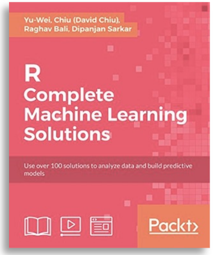


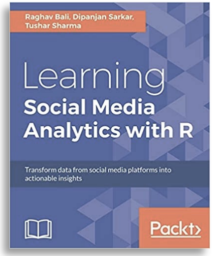
- Multiple **patents** in the field of mixed reality, deep learning, CV, NLP and Healthcare
- A number of **papers** at peer-reviewed conferences
- Well received **books** on Generative AI, NLP, Transfer Learning and more…

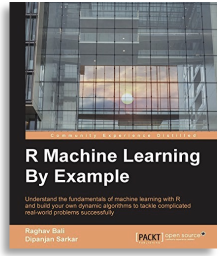










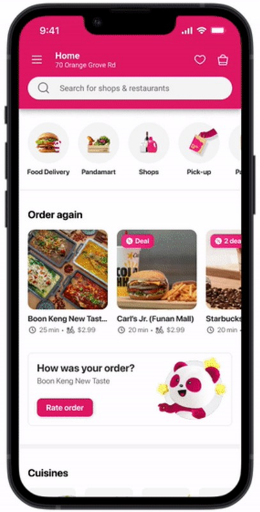


Vishal Natani

- Data Science Manager @DeliveryHero
- Experienced Data Science professional with 12+ years of experience in Embedded Systems, NLP, Computer Vision, On-Device Machine Learning solutions

Recommendation Systems

- Cold Start
- Choice Paradox
- User Geo-Diversity
- Location Preferences
- Time Preference



Data Science Workflow

EDA

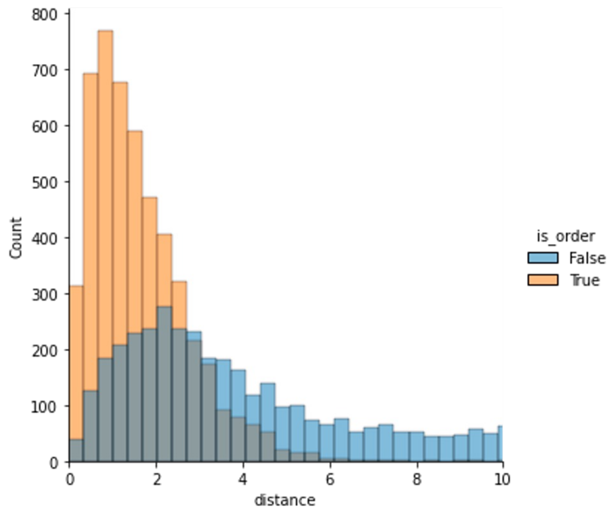
Model Building

Evaluation

Deploy

Exploratory Data Analysis

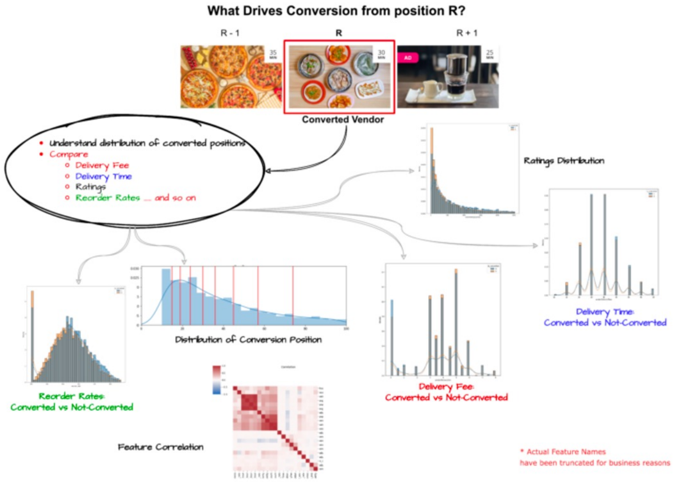
- Delivery Fee
- Popularity
- Distance
- Relevance
- More....



Distance distribution for randomly sampled ordered and non-ordered vendors. Customers prefer to order from restaurants which are closer.

Exploratory Data Analysis

- Delivery Fee
- Popularity
- Distance
- Relevance
- More....



Models

- Batch Vs Online Recommendations
- Explore / Exploit Paradigm
- Content Vs Collaborative
- …

Heuristics

- Business Rules
- Domain Knowledge
- Popularity

Linear

- Implicit features
- Matrix Factorization (SVD, NMF, ALS, SLIM)

Deep Learning

- Embeddings
- SAS-Rec, BART

LLM?

- More context and world understanding
- P5

Model Evaluation

- Different metrics help us evaluate different aspects of a recommendation engine.

MRR

Mean Reciprocal Rank

MRR is the average of the reciprocal ranks of results for a sample of queries Q

MAP

Mean Average Precision

Average precision scores at each relevant item in the ranked list of recommendations

NDCG

Normalised Discounted Cumulative Gain

Relevance of recommended items and their ranking, placing higher emphasis on the top-ranked items

More

- Recall, Coverage, Diversity, Novelty…

Business Metrics

- Offline Metrics does not translate directly into Business Metrics

CVR

Conversion Rate

CVR = (Total Conversion / Total Unique Sessions) * 100%

CTR

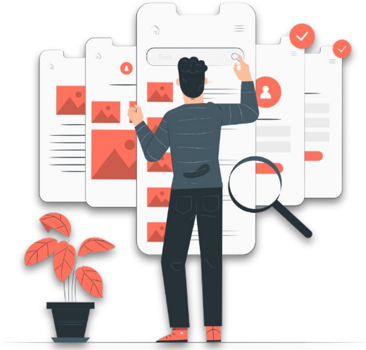
Click Through Rate

CTR = (Total Clicks / Total Impressions) * 100%

RPO

Revenue per order

RPO = (Total Revenue / Total Order) * 100%

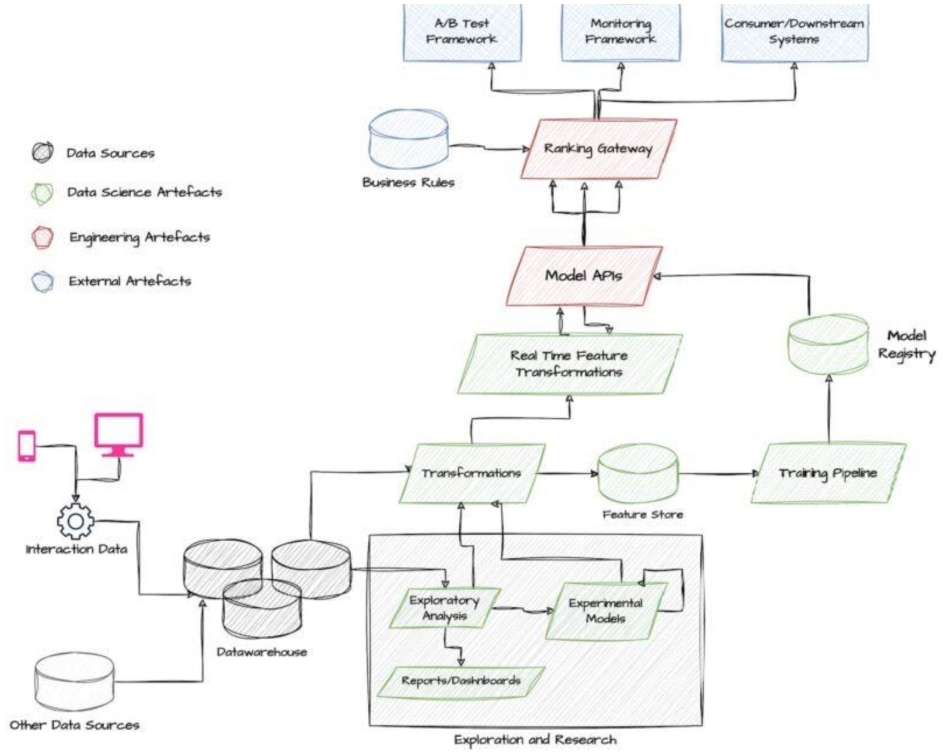


Model Performance

- **Offline Testing**: This involves using historical data to evaluate the performance of the recommendation engine, typically using metrics like precision, recall, MAP, and NDCG

- **Online Testing** (A/B Testing): In this method, two versions (A and B) of the system are compared by dividing the user base into two groups and measuring the impact on key metrics like, conversion rate, etc.

Reference Architecture

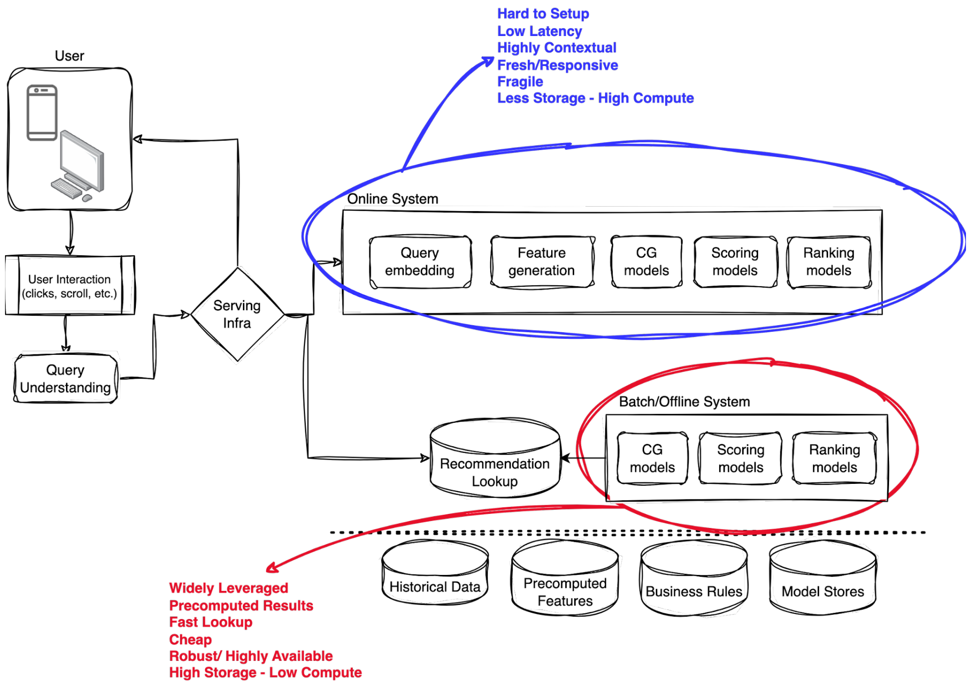


Source: https://tech.deliveryhero.com/personalisation-journey-delivery-hero/

Robust & Scalable

- **Scale**: Delivery Hero brands are serving millions of customers across hundreds of cities across the planet on a daily basis

- **Robust**: Recommendations & Ranking systems are key components for our apps. Their resilience and robustness ensure our customers get best in class experience



Open Source Dataset

Delivery Hero Recommendation Dataset

A Novel Dataset for Benchmarking Recommendation Algorithms

Paper Contributions:

- A new rich dataset to fill the gap in the online food delivery domain, enabling effective comparisons and advancements in Recommendation Systems research.

- Rich and Diverse Data: The dataset includes over a million real-world orders from three distinct cities, thousands of vendors, an extensive range of dishes, and a combined customer base of over a million individuals, ensuring a wide coverage of use cases.

City	# Orders	# Customers	# Vendors	# Products	# Products Ordered
Singapore	1.99M	512K	7411	1.06M	256K
Stockholm	400K	122K	1148	111K	41K
Taipei	2.0M	741K	9506	810K	327K

Delivery Hero Recommendation Dataset

A Novel Dataset for Benchmarking Recommendation Algorithms

Potential Use-Cases/Tasks

- Recommendations
  - What vendors to recommend for a given customer?
  - What products to recommend at checkout?
- NLP
  - How to normalize products across cuisines
  - How to recognise similar products
  - Dealing with multiple languages
- Clustering
  - What customers are similar to each other?
  - Segment customers into different behaviour groups

Scan Me To Download