Flight Rank

By: Sahar Mahar
Rehab Mohamed
Supervised by: Dr. Mona Farouk

FlightRank 2025: Aeroclub RecSys Cup | Kaggle

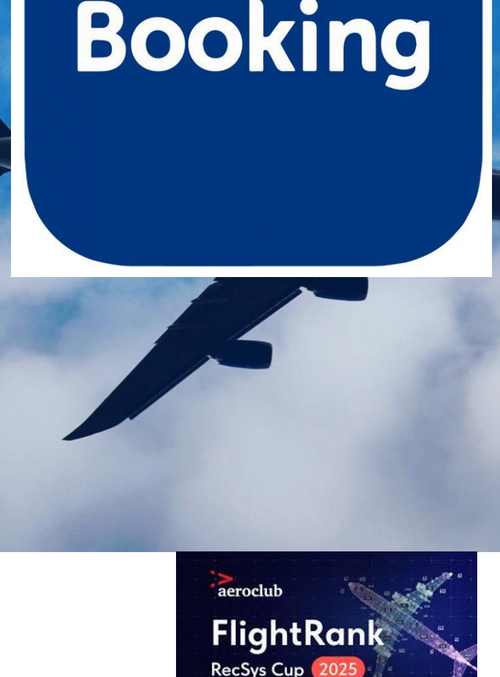
DIGITAL EGYPT BUILDERS

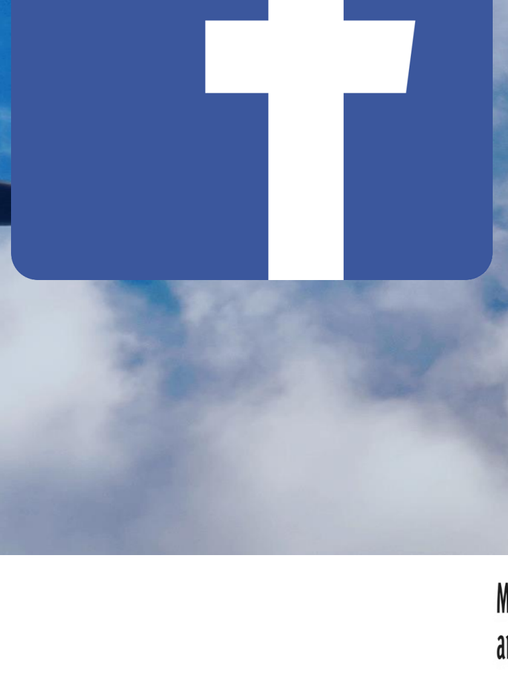
kaggle

aeroclub FlightRank RecSys Cup 2025

Ministry of Communications and Information Technology

How Many of you uses YouTube, Facebook, Booking, etc.. Apps?



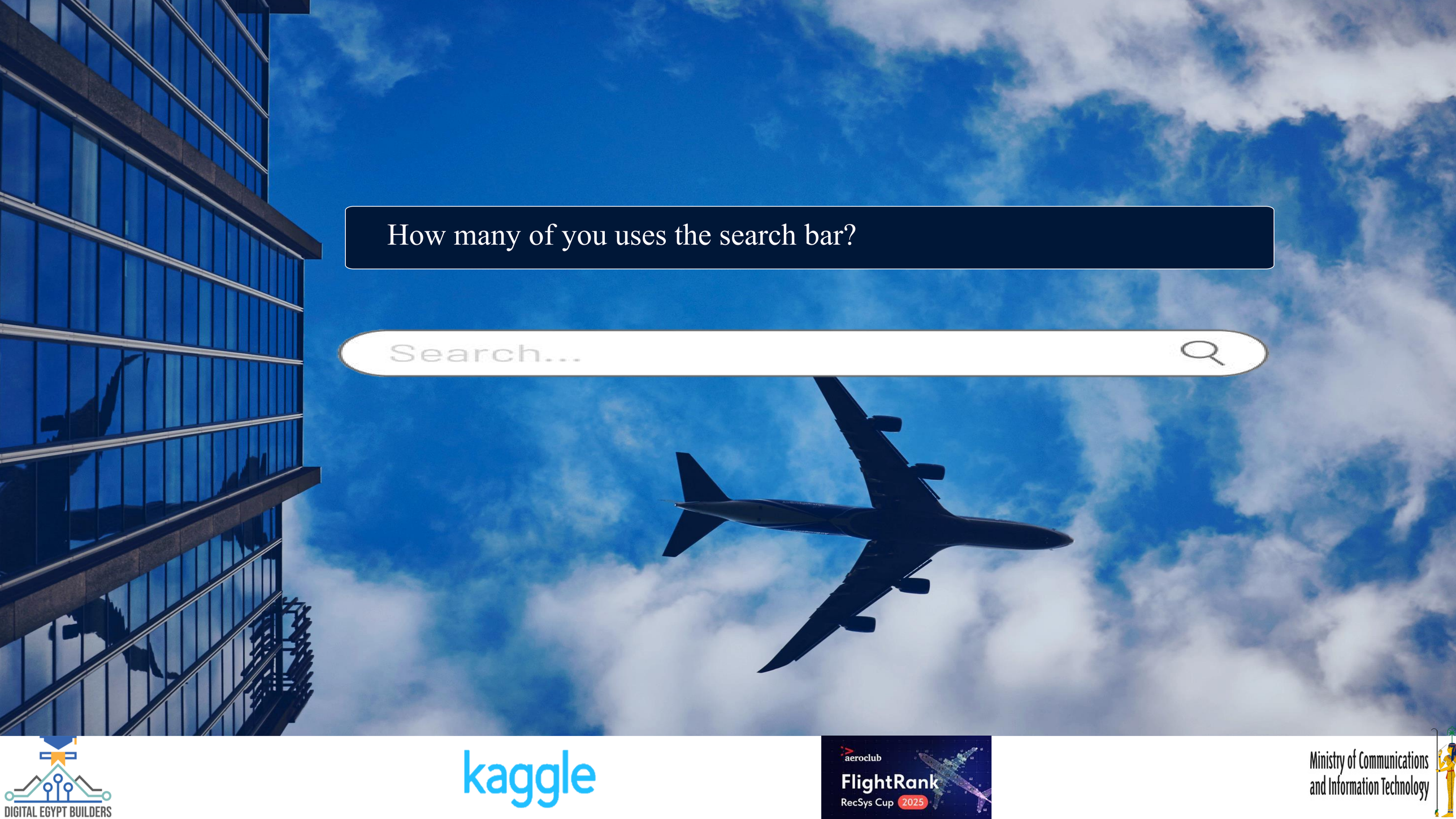


DIGITAL EGYPT BUILDERS

kaggle

aeroclub
FlightRank
RecSys Cup 2025

Ministry of Communications and Information Technology



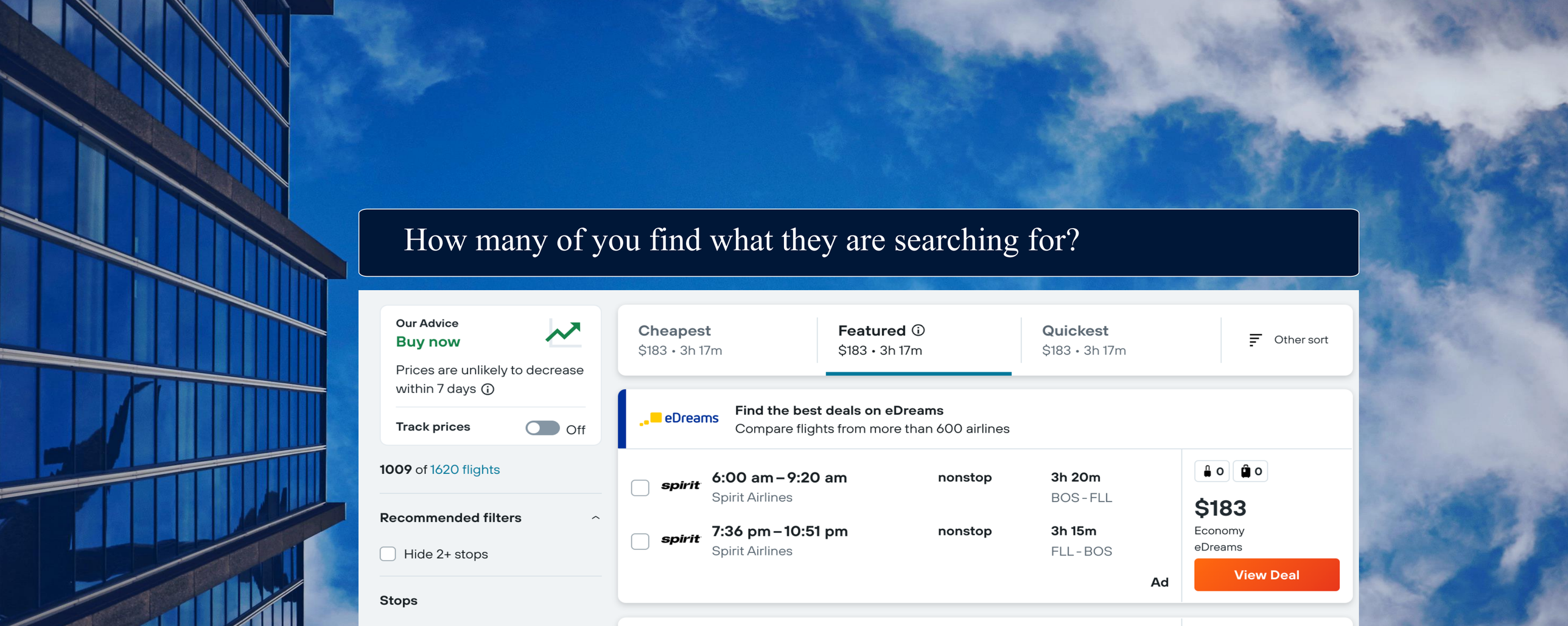
DIGITAL EGYPT BUILDERS

kaggle

aeroclub
FlightRank
RecSys Cup 2025

Ministry of Communications and Information Technology

How many of you find what they are searching for?



DIGITAL EGYPT BUILDERS

kaggle

aeroclub
FlightRank
RecSys Cup 2025

Ministry of Communications and Information Technology

Table Of Content

Problem Definition

- Build a recommendation system that predicts which flight option a business traveler is most likely to choose within a given search session.
- The model has to rank all the available flight options from best to worst based on the user's specific preferences, company policies and other factors.

DIGITAL EGYPT BUILDERS

kaggle

aeroclub FlightRank RecSys Cup 2025

Ministry of Communications and Information Technology

Motivation

- Our goal is to create a model that predicts future airline recommendation system's performance using a large and complex. dataset from a global competition.
- Learn how to deal with real-world problem.

DIGITAL EGYPT BUILDERS

kaggle

aeroclub FlightRank RecSys Cup 2025

Ministry of Communications and Information Technology

Methodology

Data

- Each search session = ranker_id → multiple flight options
- Train: selected (1 = chosen, 0 = not chosen)
- Test: predict rank (1= Best, N = worst)
- train.parquet → Training data
- Test.parquet → Test data (no target)

Data

- User Info: gender, nationality, frequent flyer, VIP
- Company Policy: tariff code, policy compliance
- Route & Search: route, request date/time
- Pricing: total price, taxes
- Flight Timing: outbound/return times & durations
- Segments: airline, aircraft, baggage, seats, cabin class
- Rules: cancellation & exchange penalties

Data Analysis

- Drop columns with null values > 98%
- Drop unnecessary columns.
- Convert data type columns from string into date & time using regex and polar functions.
- Total Price

Model Selection

XGBoost

1. Performance (Accuracy / Ranking Score)

- 38%

2. Training & Inference Speed

- Fast training, but can be slower with large categorical data (needs one hot encoding).

3. Handling of Categorical Features

- Needs preprocessing (label encoding or one-hot).

4. Robustness & Ease of Use

- Requires more manual hyper parameter tuning to reach good performance.

CatBoost

1. Performance (Accuracy / Ranking Score)

- 40%

2. Training & Inference Speed

- Automatically handles categorical features efficiently, usually faster on inference when many categorical features exist.

3. Handling of Categorical Features

- Native categorical handling no need for heavy preprocessing.

4. Robustness & Ease of Use

- Often works well out of the box with fewer parameter.

Challenges

1. Data size = 5.27 GB
2. CPU & TPU VM v3-8 limitation
3. Selecting the best model


Demo

Result

FlightRank 2025: Aeroclub RecSys Cup

Late Submission ...

Overview | Data | Code | Models | Discussion | Leaderboard | Rules | Team | Submissions

#	Team	Members	Score	Entries	Last	Solution
1	gezi		0.54062	100	3d	
2	Phaedrus		0.53548	51	3d	
3	AF		0.53548	190	3d	
4	mango789 & Ionut Visan		0.53438	210	3d	
5	Taos David		0.52970	56	3d	
6	Mikhail Golubchik		0.52814	105	3d	
7	kif		0.52795	63	4d	
8	XS 330		0.52529	39	3d	
9	Muhammad Hassaan		0.52446	156	3d	
10	Dmitriy Ch		0.52437	134	4d	

DIGITAL EGYPT BUILDERS

kaggle

aeroclub FlightRank RecSys Cup 2025

Ministry of Communications and Information Technology

Result

FlightRank 2025: Aeroclub RecSys Cup

Late Submission

Overview | Data | Code | Models | Discussion | Leaderboard | Rules | Team | Submissions

437	▼ 2	Sagor Kumar Mitra	0.40540	2	3d
438	▼ 4	Dayou Wang	0.40459	1	7d
439	▼ 2	Jay Prajapati	0.40451	1	1mo
440	—	Zhong Haitian	0.40354	7	2mo
441	—	akmonymous	0.40350	2	1mo
442	—	sahar mah3r99	0.40226	10	2d

DIGITAL EGYPT BUILDERS

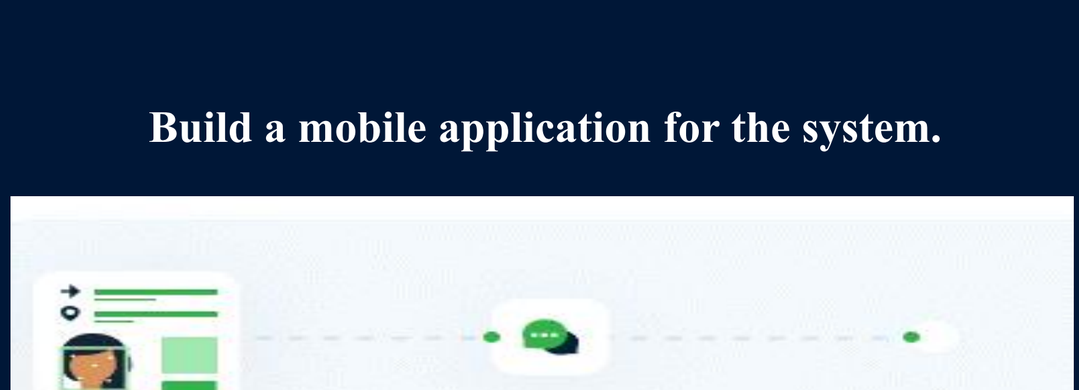
kaggle

aeroclub FlightRank RecSys Cup 2025

Ministry of Communications and Information Technology

Future Work

Build a mobile application for the system.



DIGITAL EGYPT BUILDERS

kaggle

aeroclub
FlightRank
RecSys Cup 2025

Ministry of Communications
and Information Technology