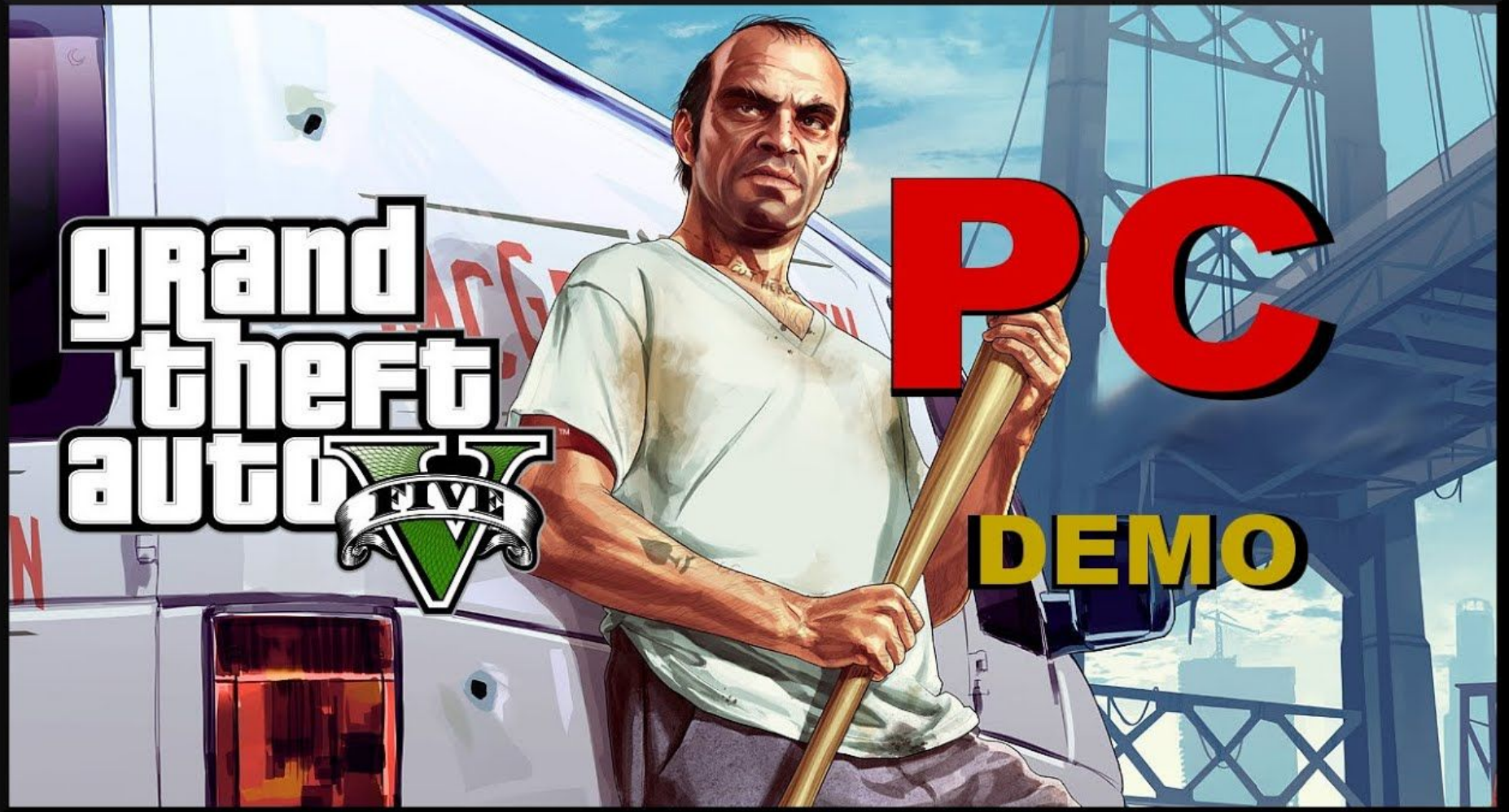
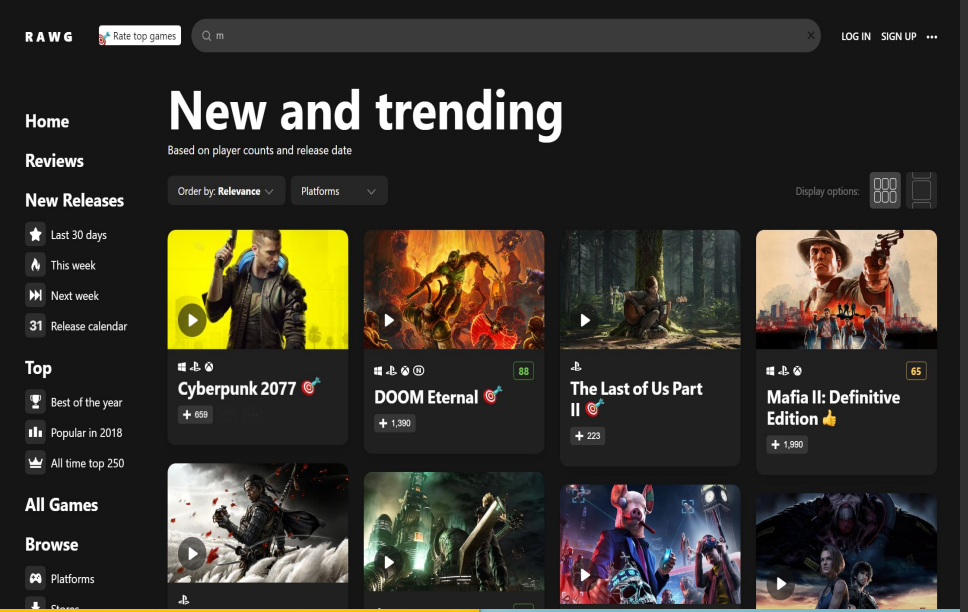
THE GAMEXPLORER

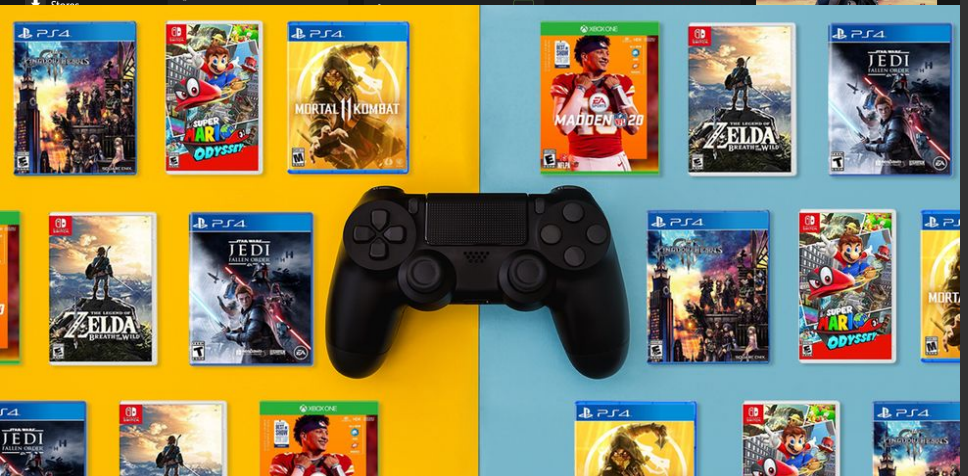
Bringing your inner gamer out, this quarantine

ADITYA GUDAL

The Problem: Reliving my gaming days during quarantine, but in moderation.


grand theft auto
FIVE
PC
DEMO





Data

- RAWG database - ~350,000 video games over 50 platforms.

  API docs link:https://rawg.io/apidocs

- API fetches 3240 games with their characteristics like playtime, user score, rating, genres, review counts, suggestion counts etc.

- The RDS instance stores the model's 10 recommendations and the user's favorite game

Model and Success

1. Models:
   a. K-nearest neighbors - not bad
   b. K-means clustering - not bad
   c. Alternating Least Squares - better

Best Hyperparameters: 25 latent factors, 0.1 shrinkage, 100 iterations

2. Success Criteria:
   a. Machine Learning: Root Mean Square Error < 0.4
   b. Business: Atleast 50,000 User sessions and positive feedback count > 3000.
3. Results:
   a. Root Mean square error - 0.49
   b. Mean absolute error - 0.44

ALS:https://blog.insightdatascience.com/explicit-matrix-factorization-als-sgd-and-all-that-jazz-b00e4d9b21ea
Data Split:https://nbviewer.jupyter.org/github/jmsteinw/Notebooks/blob/master/RecEngine_NB.ipynb

Matrix factorization - ALS

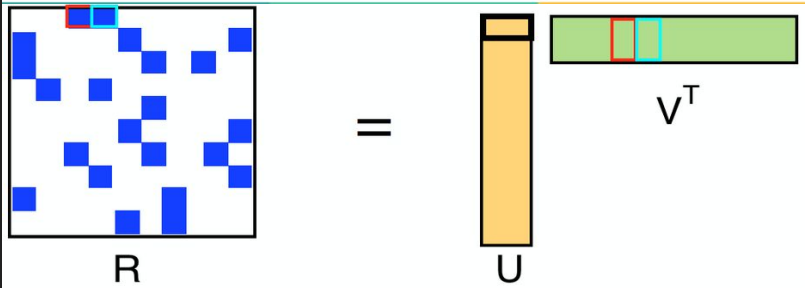
- Fix P and optimize Q
- Convex problem when one of the latent matrices is fixed
- Easy to parallelize

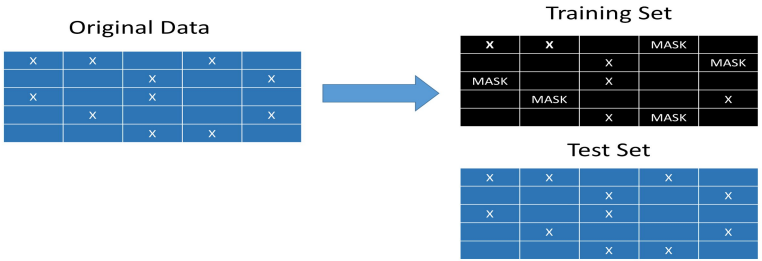
Losses:

$$\forall p_i : L(p_i) = ||R_i - P_i \times Q^T||_2 + \lambda \cdot ||p_i||_2$$
$$\forall q_j : L(q_j) = ||R_i - P \times Q_j^T||_2 + \lambda \cdot ||q_j||_2$$

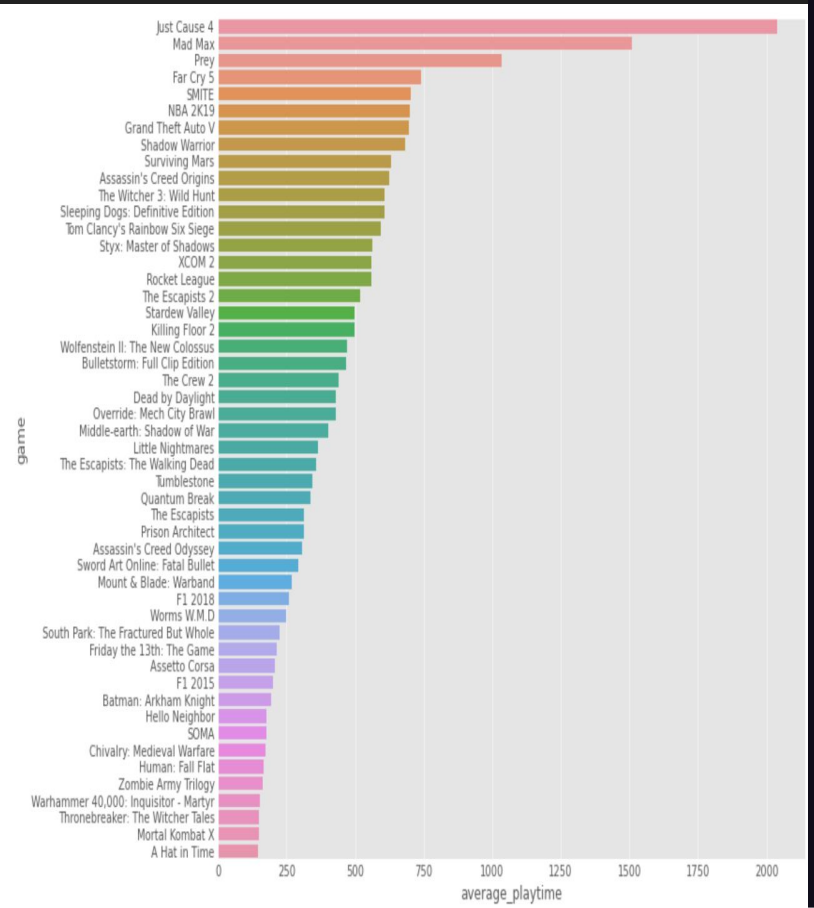
Solutions:

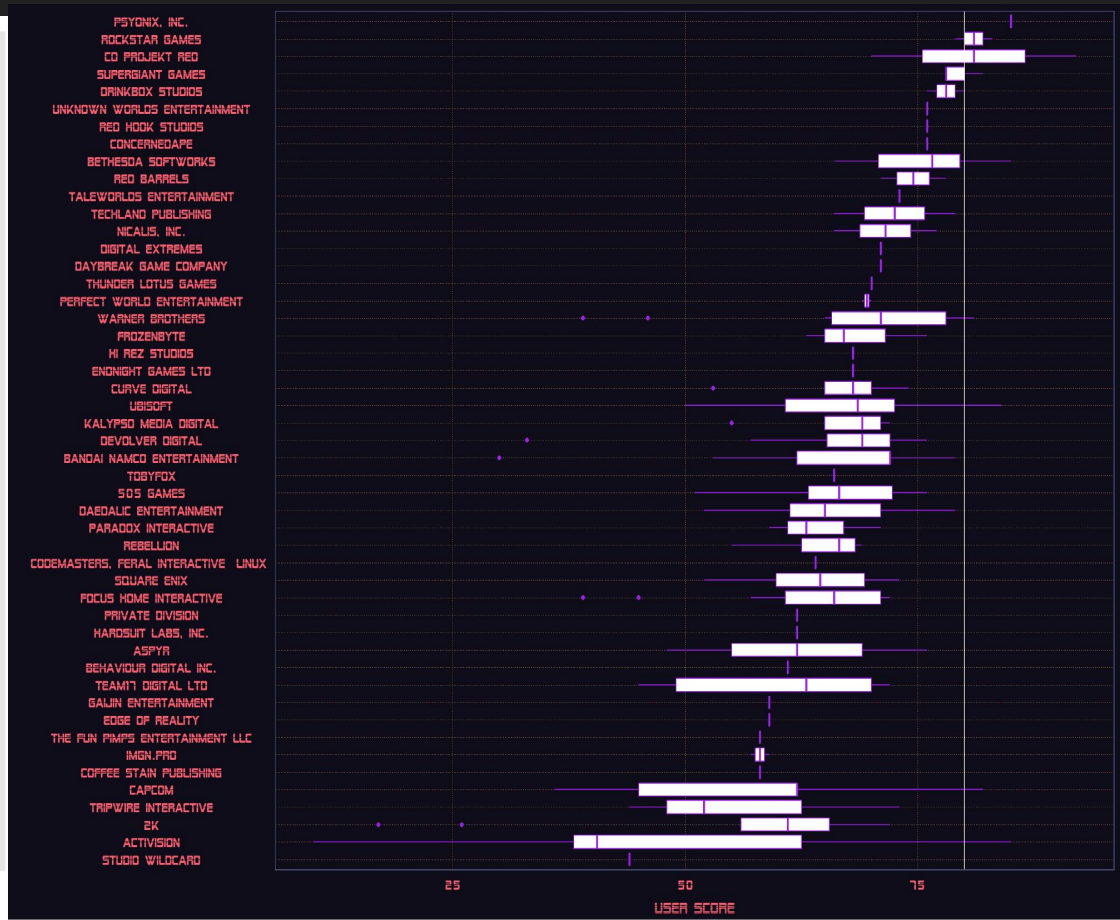
$$p_i = \left(Q^T \times Q + \lambda I\right)^{-1} \times Q^T \times R_i$$
$$q_j = \left(P^T \times P + \lambda I\right)^{-1} \times P^T \times R_j$$





Interesting Insights





University of Mumbai

Github: https://github.com/aditya9729

Linkedin: https://www.linkedin.com/in/aditya-gudal

Email: adityagudal2020@u.northwestern.edu

Northwestern University