Computer Engineering Department

G33 Travel Destination Recommender System

Project Advisor: Prof. Magdalini Eirinaki

Devineni, Geethika (MS Software Engineering)
Garlapati, Arun (MS Software Engineering)
Kankanala, Venkatesh (MS Software Engineering)
Mungi, Sumanth Reddy (MS Artificial Intelligence)

Introduction

Welcome to our innovative travel destination recommender system! We take the stress out of trip planning by using cutting-edge technology to create the perfect itinerary for you.

- No more hours spent scouring websites and guidebooks.
- Personalized recommendations based on your interests, budget, and travel style.
- Discover hidden gems and unique experiences you might have missed.
- Make informed decisions with up-to-date travel data at your fingertips.

Methodology

Alternating Least Squares - Collaborative filtering

- The system gathers information about the user's preferences through explicit user input such as ratings and implicit user input such as behavior. This can include past trips, searches for destinations with specific features, or browsing history.

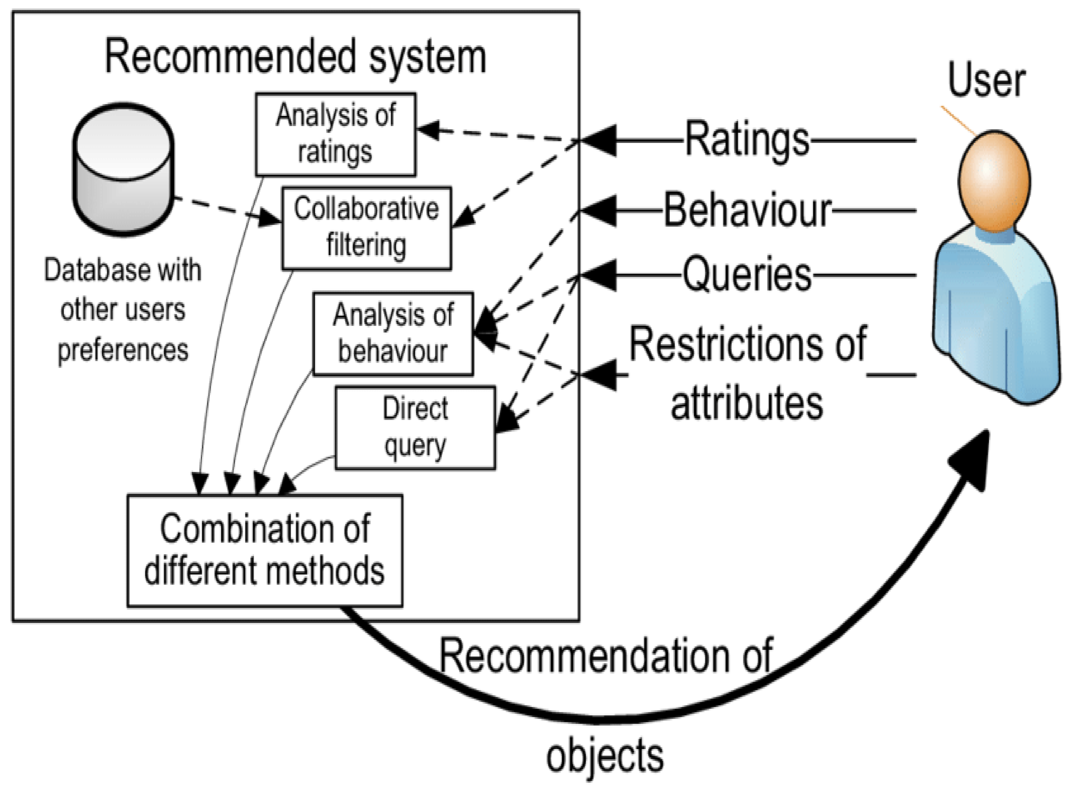
- This method combines collaborative filtering and content-based filtering to make recommendations. This is a popular approach because it can address the limitations of each individual method.

- $R \approx U \times V$ user's historical interactions with hotels and restaurants can be approximated by multiplying two matrices – U (user's latent preferences) and V (hotels and restaurants' latent features).

$$\hat{R}_{ij} = \sum_{k=1}^{K} U_{ik} \times V_{kj}$$

- Where R (i, j) is the predicted interaction between user i and item j, U (i, k) represents user i's preference for k-th feature and V (k ,j) represents j-th item's association with k-th latent feature.

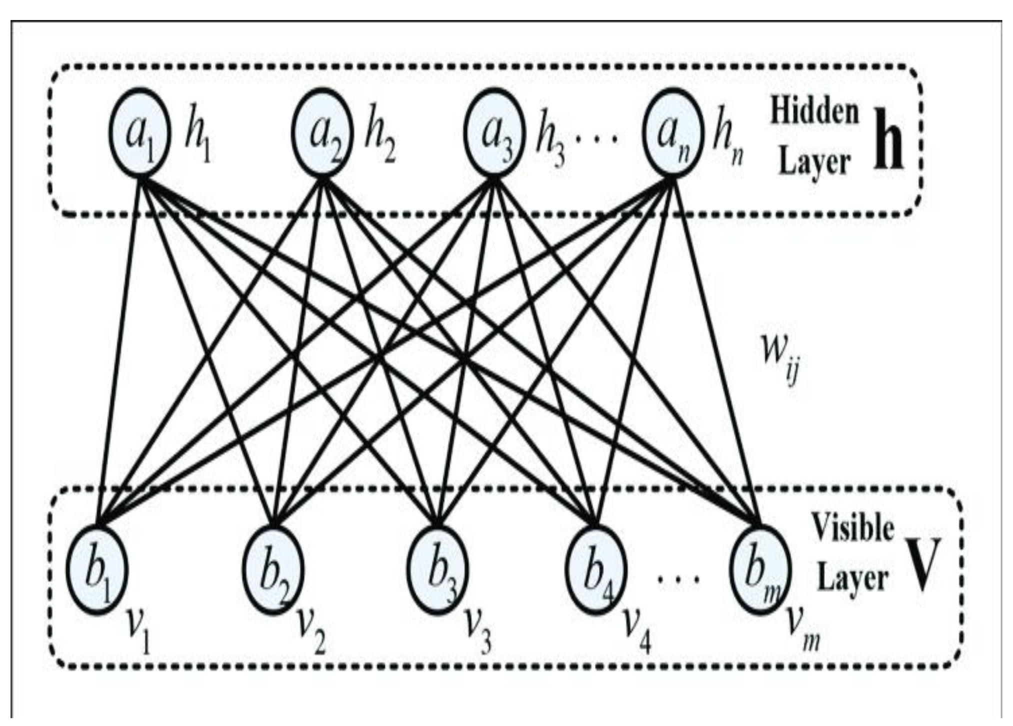
Methodology



- The matrices U and V are learned iteratively using optimization techniques such as ALS, minimizing the reconstruction error between the predicted and observed interactions.

- This factorization enables us to discover latent features that explain user-item interactions, facilitating personalized recommendations based on these underlying patterns and preferences in user behavior.

Restricted Boltzmann Machine Model



- This bottom layer represents the input data. Each circle (v1, v2, ..., vm) denotes a visible unit that corresponds to a single feature in the input data. The top layer in the diagram represents the hidden features learned by the RBM. Each circle (h1, h2, ..., hn) denotes a hidden unit that captures higher-level patterns from the input data.

- In our recommender system, these features represent user preferences like preferred location, budget, and activity type.

- The lines connecting the circles represent connections (weights) between the visible and hidden units.

- The RBM uses these connections and unit activations to learn a probability distribution over the input data. This allows it to perform tasks like dimensionality reduction, feature learning, and data reconstruction.

- The formula explanation for RBM is as follows:

- Energy Function (E): evaluates compatibility between attractions (visible units v) and user preferences (hidden units h). RBM calculates the energy associated with the features of attractions (*ai vi)* and user preferences (bj hj) – helping in identifying user's interests and attractions that align with their travel profile.

$$E(v,h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i h_j w_{ij}$$

- Probability of hidden unit activation(P(h/v)): calculates the probability of hidden units(h) being activated given visible units values (v). Uses sigmoid function.

$$P(h_j = 1|v) = \sigma\left(b_j + \sum_i v_i w_{ij}\right)$$

- Probability of visible unit activation(P(v/h)): calculates the probability of visible units(v) being activated given hidden units values (h). Uses sigmoid function.

$$P(v_i = 1|h) = \sigma\left(a_i + \sum_j h_j w_{ij}\right)$$

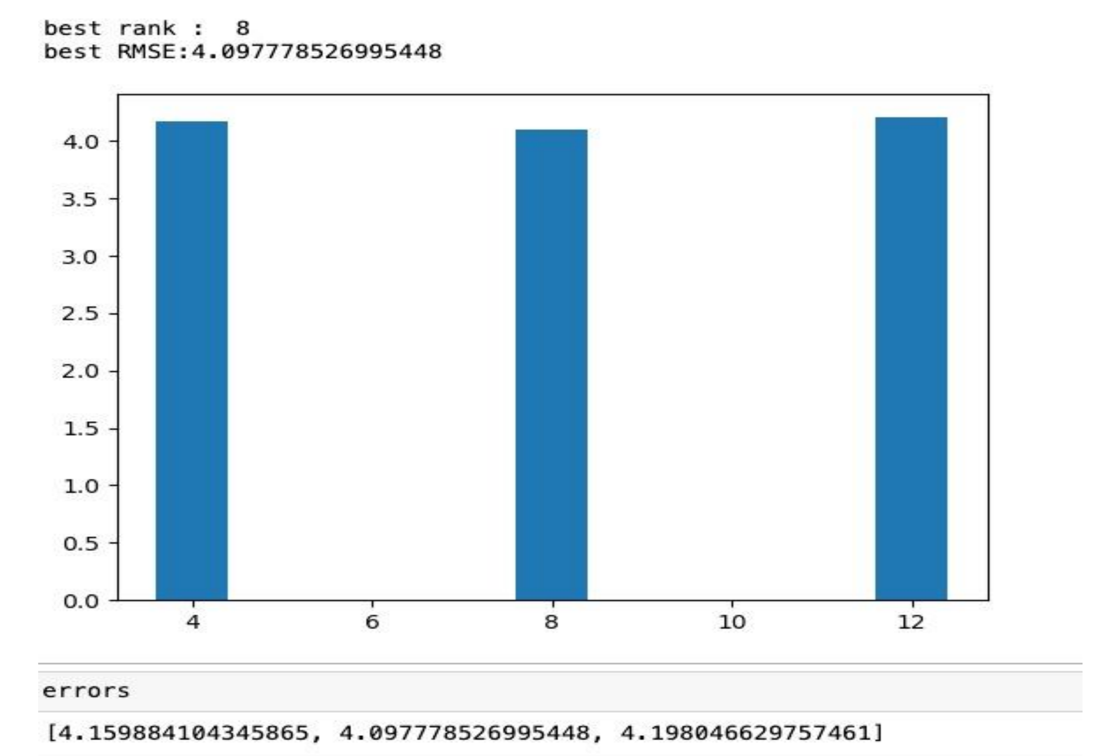
- Contrastive Divergence (CD-k): used to approximate the gradient of the log-likelihood function in training RBMs. Consists of multiple Gibbs sampling to estimate the difference between observed data and data generated by RBM. Iteratively adjusts weights and biases of RBM to minimize the difference.

Analysis and Results

- The graph covers a 12-month period, with the x-axis labeled "months" ranging from 4 to 12, and the y-axis labeled "errors" ranging from 0 to 4.1 errors.

- The number of errors appears to be relatively stable over the past 12 months, with some slight fluctuations.

- The first data point is at month 4, with a value of approximately 4.16 errors. The second data point is at month 8, with a value of approximately 4.1 errors. The third data point is at month 12, with a value of approximately 4.2 errors.



- The best RMSE (root mean squared error) is 4.0977, which is a measure of the difference between the predicted values and the actual values. A lower RMSE indicates a better fit for the model.

Summary/Conclusions

- Recommends destinations based on individual preferences, budget, and travel timing, empowering informed decisions and new discoveries.
- Cloud infrastructure, data processing pipelines, and advanced machine learning algorithms ensure efficiency, scalability, and personalization.
- This travel destination recommender system stands as a game-changer, streamlining trip planning and enhancing the overall travel experience globally.

Key References

[1] H. Ko, S. Lee, Y. Park, and A. Choi, "A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields, "in *Electronics*, vol. 11, no. 1, pp. 141, 2022.

[2] U. Gretzel, N. Mitsche, Y. H. Hwang, and D. R. Fesenmaier, "Tell Me Who You Are, and I Will Tell You Where to Go: Use of Travel Personalities in Destination Recommendation Systems," in *Information Technology & Tourism*, vol. 7, no. 1, pp. 3–12, 2004.

[3] M. Goossen, H. Meeuwsen, J. Franke, and M. Kuyper, "My Ideal Tourism Destination: Personalized Destination Recommendation System Combining Individual Preferences and GIS Data," in *Information Technology & Tourism*, vol. 11, no. 1, pp. 17-30, 2009.

Acknowledgements

The authors are deeply indebted to Prof. Magdalini Eirinaki