Recommender System : Collaborative Filtering

Nan An
Othman Hicheur
Kanupriya Jain

October 2, 2024

Table of Contents

Overview (1/2)

METHODS USED:

- **Collaborative Filtering (CF)**: Leverages user-item interactions to make recommendations.
  - **ALS (Matrix Factorization)**: Decomposes user-item data into latent factors.
  - **NCF (Neural Collaborative Filtering)**: Uses neural network for better flexibility.
  - **Genre-Enhanced NCF**: Adds content-based features (like genre) for better predictions.

Overview(2/2)

- **Probabilistic Collaborative Filtering (PCF)**: Models uncertainty in user preferences, enhancing prediction accuracy by using a probabilistic approach.
- **Post-Processing**: Applied across all methods to refine rankings, improve diversity, and increase relevance.

Table of Contents

2 ALS Matrix Factorization

ALS Matrix Factorization

- **Matrix Factorization:**

$$R \approx U \times I^T$$

- **Loss Function:**

$$\mathcal{L}(U, I) = \sum_{(u,i)\in\mathcal{K}} (r_{ui} - U_u \cdot I_i)^2 + \lambda(||U_u||^2 + ||I_i||^2)$$

where:

— r_{ui} is the actual rating given by user u for item i.
— $\mathcal{K}$ is the set of observed ratings.
— λ is the regularization parameter to prevent overfitting.

ALS Matrix Factorization (2/2)

Library Used:

- `implicit`

Best Hyperparameters

- **Factors:** 10
- **Regularization:** 0.01
- **Iterations:** 50

Performance Metrics

- **Mean Squared Error (MSE):** 12.05
- **Root Mean Squared Error (RMSE):** 3.47

Table of Contents

3 Neural Collaborative Filtering

Cross-Validation for Selecting neurons Combination

Splitting Training Data(5 folds)

↓

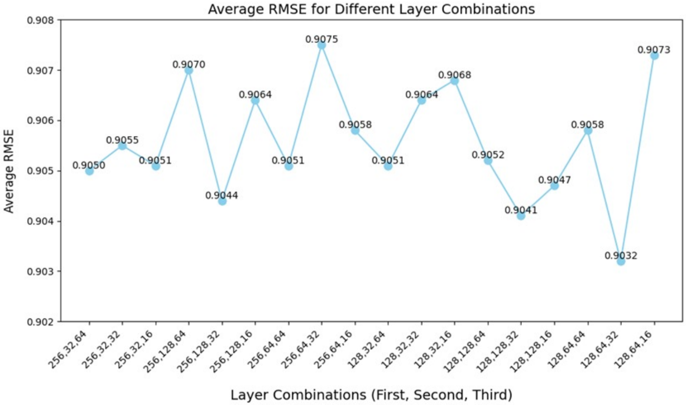
Building Dense Layer neurons

↓

Averaging the Results

↓

Selecting the Best Combination



Decision Diagram for building the NCF

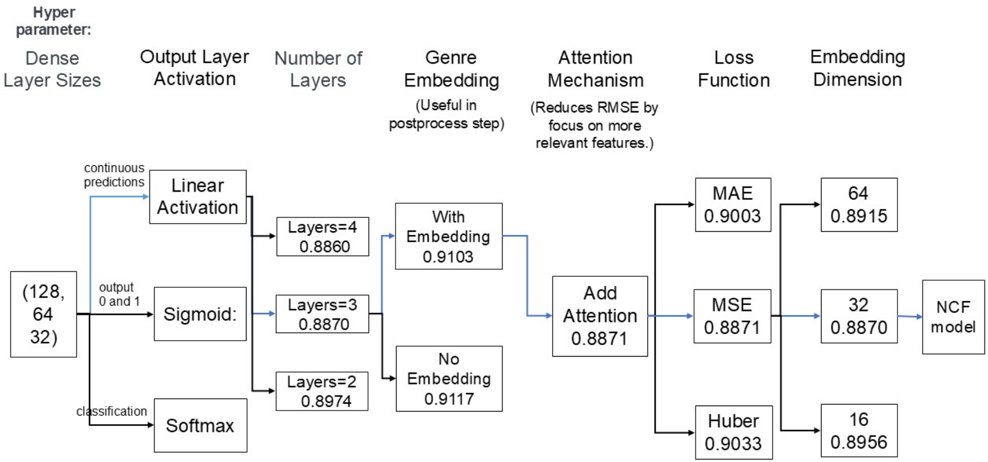


Table of Contents

Neural Collaborative Filtering with genre and attention layer

Key Concepts:

- **Training with genre:** Same as before, but add of one-hot encoding representing the genre of movie, in order to catch the relationship between genre and user for each one.
- **Attention layer:** Use of attention layer in order to put "weights" on some genre of movie for each user.

Results:

`embeddings=32`

RMSE without hyper parametrization and post-processing = 0.861

Table of Contents

Post Processing

Key Concepts:

- **Rounding to a semi-integer:** Classifier as we know there are only 9 ratings, put to the nearest semi-integer, use of distribution of ratings and Bayes prediction.
- **Use of distribution and residuals:** Plot the residuals over the real ratings, see how much the predictions are far from the reality and adjust from this.

Table of Contents

6 Probabilistic Matrix Factorization

Probabilistic Matrix Factorization(1/3)

Key Concepts:

- **Gaussian Assumption:**
  - Each observed rating r_{ij} is modeled as a Gaussian:

$$r_{ij} \sim \mathcal{N}(u_i^T v_j, \sigma^2)$$

- **Log-Likelihood:**
  - The likelihood of observed ratings given user and item latent factors:

$$\log p(R|U, V, \sigma^2) = -\frac{1}{2\sigma^2} \sum_{(i,j)\in\text{observed}} (r_{ij} - u_i^T v_j)^2$$

Probabilistic Matrix Factorization(2/3)

Key Concepts:

- **MLE (Maximum Likelihood Estimation):**
  - Maximize the log-likelihood to estimate the latent factors U and V.
  - Equivalent to minimizing the Mean Squared Error (MSE).
- **Regularization (Priors):**
  - Gaussian priors on latent vectors $u_i \sim \mathcal{N}(0, \lambda^{-1} I)$ and $v_j \sim \mathcal{N}(0, \lambda^{-1} I)$ help prevent overfitting:

$$\log p(U, V | R, \sigma^2) = \text{MSE} + \lambda \left(\|u_i\|^2 + \|v_j\|^2 \right)$$

Probabilistic Matrix Factorization(3/3)

Results:

$num\_factors = 10, epochs = 1000$

- **Withput Post Processing:**
  RMSE without hyperparameter tuning = 1.286
- **With Post Processing:**
  RMSE without hyperparameter tuning = 1.156

Table of Contents

Future Work

- Post-Processing
- Optimizing the Attention Mechanism
  - Dynamic Attention Weights: Investigate weights dynamic assigned to users, items, and genres adapt based on context
  - Multi-Head Attention
  - Attention Regularization

Q&A

Thank you for listening!