Shrey Shah
CWID - 20009523

Recommender Systems

CS 541 B - Artificial intelligence

A recommender system, or a recommendation system (sometimes replacing 'system' with a synonym such as platform or engine), is a subclass of information filtering system that provide suggestions for items that are most pertinent to a particular user. Typically, the suggestions refer to various decision-making processes, such as what product to purchase, what music to listen to, or what online news to read. Recommender systems are particularly useful when an individual needs to choose an item from a potentially overwhelming number of items that a service may offer.

Recommender systems are used in a variety of areas, with commonly recognised examples taking the form of playlist generators for video and music services, product recommenders for online stores, or content recommenders for social media platforms and open web content recommenders. These systems can operate using a single input, like music, or multiple inputs within and across platforms like news, books and search queries. There are also popular recommender systems for specific topics like restaurants and online dating. Recommender systems have also been developed to explore research articles and experts, collaborators, and financial services.

Dataset:

This dataset contains a set of movie ratings from the MovieLens website, a movie recommendation service. This dataset was collected and maintained by GroupLens, a research group at the University of Minnesota. There are 5 versions included: "25m", "latest-small", "100k", "1m", "20m". In all datasets, the movies data and ratings data are joined on "movieId". The 25m dataset, latest-small dataset, and 20m dataset contain only movie data and rating data. The 1m dataset and 100k dataset contain demographic data in addition to movie and rating data.

- "25m": This is the latest stable version of the MovieLens dataset. It is recommended for research purposes.
- "latest-small": This is a small subset of the latest version of the MovieLens dataset. It is changed and updated over time by GroupLens.
- "100k": This is the oldest version of the MovieLens datasets. It is a small dataset with demographic data.
- "1m": This is the largest MovieLens dataset that contains demographic data.
- "20m": This is one of the most used MovieLens datasets in academic papers along with the 1m dataset.

For each version, users can view either only the movies data by adding the "-movies" suffix (e.g. "25m-movies") or the ratings data joined with the movies data (and users data in the 1m and 100k datasets) by adding the "-ratings" suffix (e.g. "25m-ratings").

The features below are included in all versions with the "-ratings" suffix.

- "movie_id": a unique identifier of the rated movie
- "movie_title": the title of the rated movie with the release year in parentheses

- "movie_genres": a sequence of genres to which the rated movie belongs
- "user_id": a unique identifier of the user who made the rating
- "user_rating": the score of the rating on a five-star scale
- "timestamp": the timestamp of the ratings, represented in seconds since midnight Coordinated Universal Time (UTC) of January 1, 1970

Learning:

After we terminate SGD, we will obtain the solution U, V . Our prediction matrix X is then given by X = UV^t. We evaluate the performance of our prediction matrix X by root-mean-square error (RMSE):

$$\min_{U,V} F(U,V) := \frac{1}{2} \sum_{(i,j)\in\Omega_1} (M_{ij} - \boldsymbol{u}_i \boldsymbol{v}_j^\top)^2 + \frac{\lambda}{2}(\|U\|_F^2 + \|V\|_F^2)$$

where Mij is the (i, j)th entry of M, ui and vj are the ith and jth row of U and V respectively.

1. For a given index (i, j), derive the stochastic gradient $\frac{\partial F(U,V)}{\partial \boldsymbol{u}_i}$ and $\frac{\partial F(U,V)}{\partial \boldsymbol{v}_i}$.

2. Suppose = 1. Describe the update rule of SGD and implement it with Python. You can randomly initialize all u_i and v_j . Note that you need to carefully choose the learning rate.

3. Plot the objective value against the number of iterations, and summarize your findings.

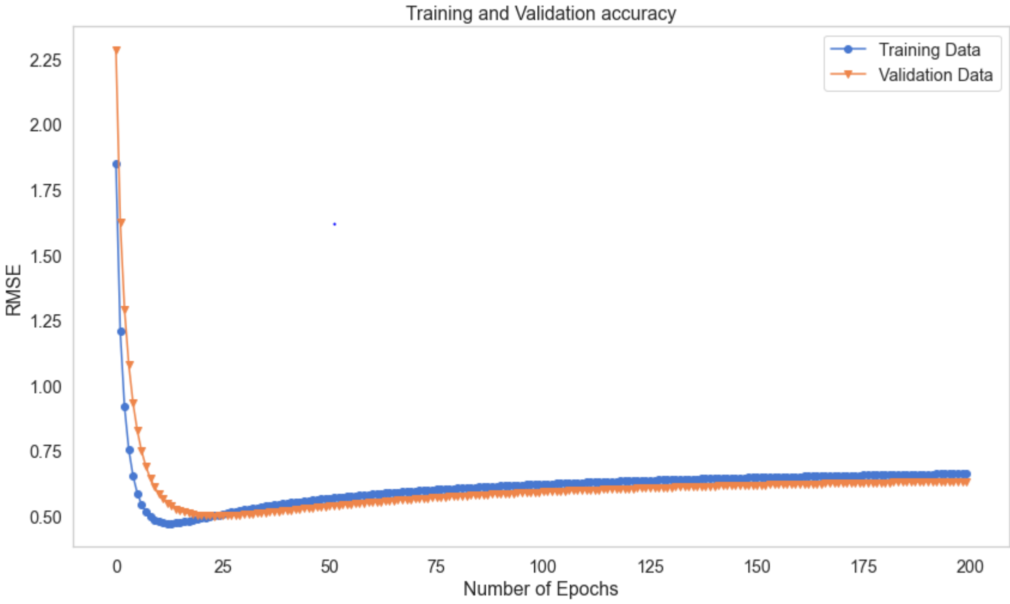
Evaluation:

After we terminate SGD, we will obtain the solution U, V . Our prediction matrix X is then given by X = UV^t. We evaluate the performance of our prediction matrix X by root-mean-square error (RMSE):

$$\text{RMSE} := \sqrt{\frac{1}{|\Omega_2|} \sum_{(i,j)\in\Omega_2} (M_{ij} - X_{ij})^2}.$$

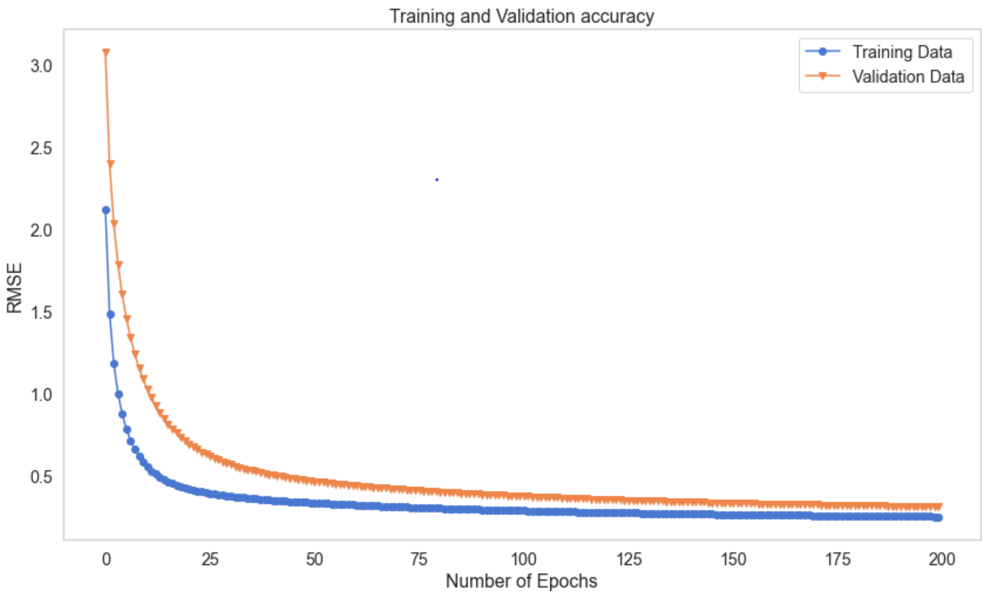
Shrey Shah
CWID - 20009523

1. Record the RMSE for the choice **λ** = 1.



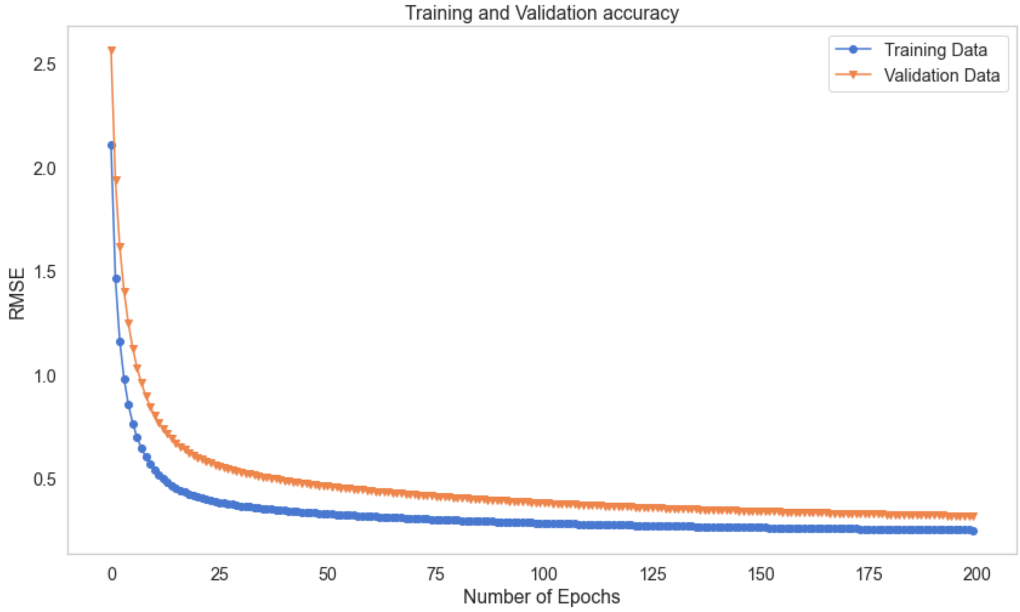
2. Now pick from {10^{-6}, 10^{-3}, 0.1, 0.5, 2, 5, 10, 20, 50, 100, 500, 1000}. For each value, learn and evaluate the your model. Plot RMSE against and summarize your findings.

λ = 0.000001

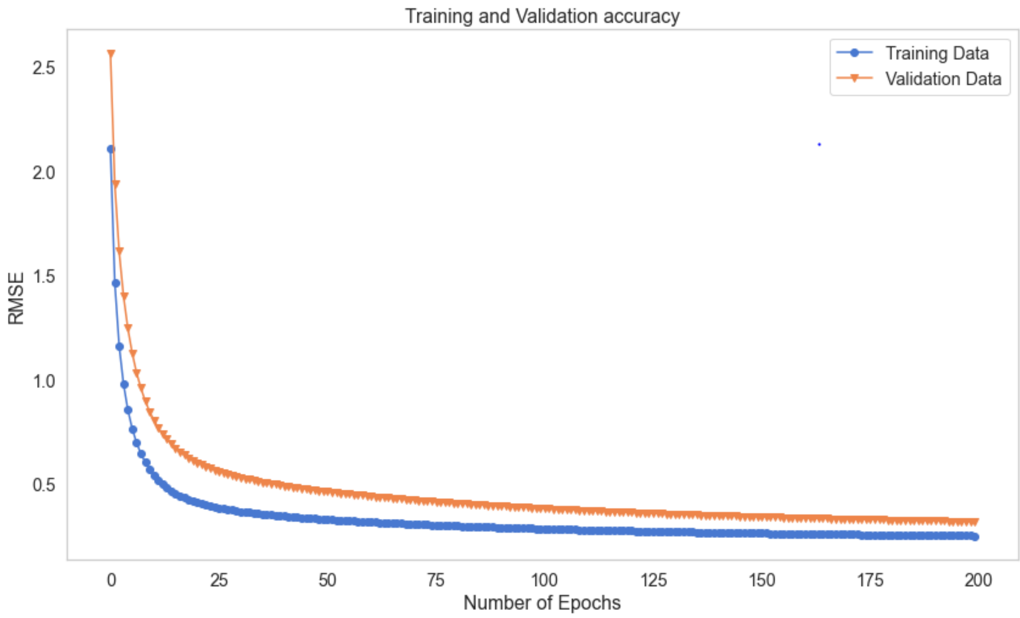


Shrey Shah
CWID - 20009523

λ = 0.003

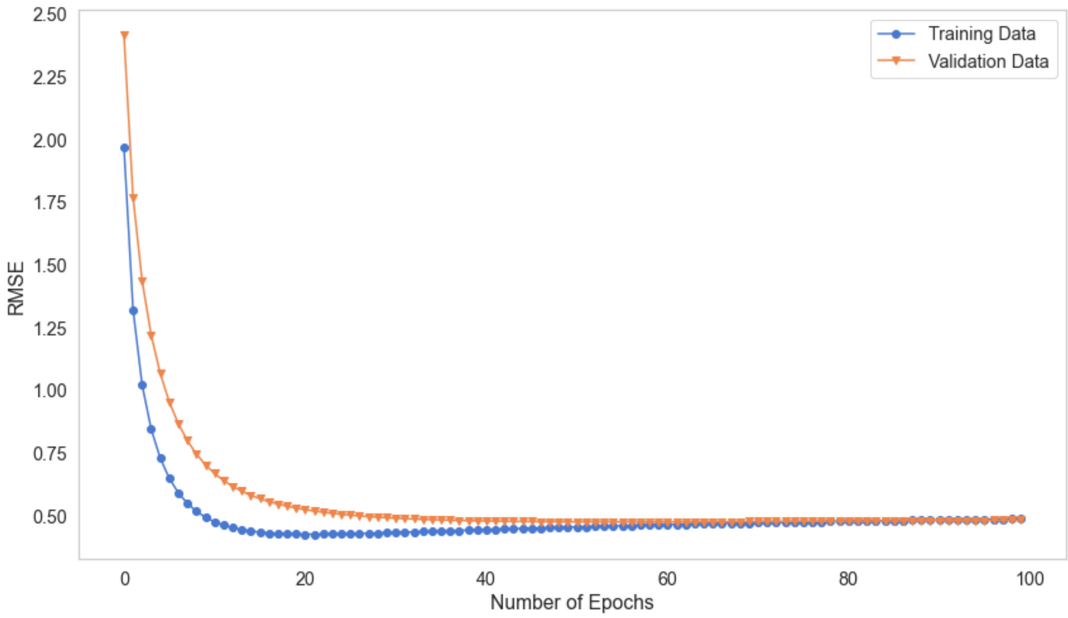


λ = 0.1



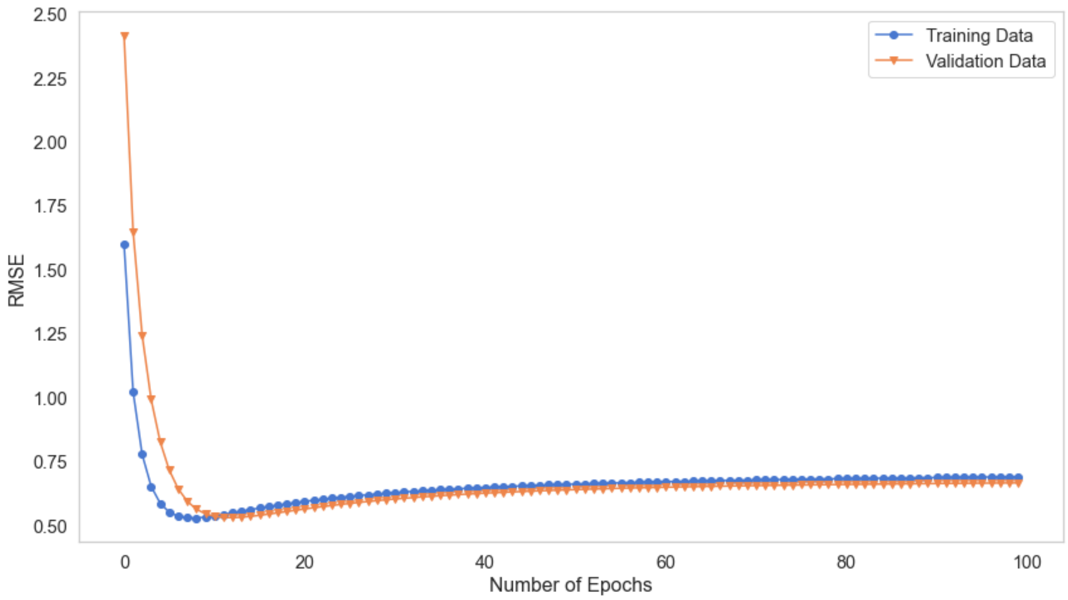
Shrey Shah
CWID - 20009523

λ = 0.5



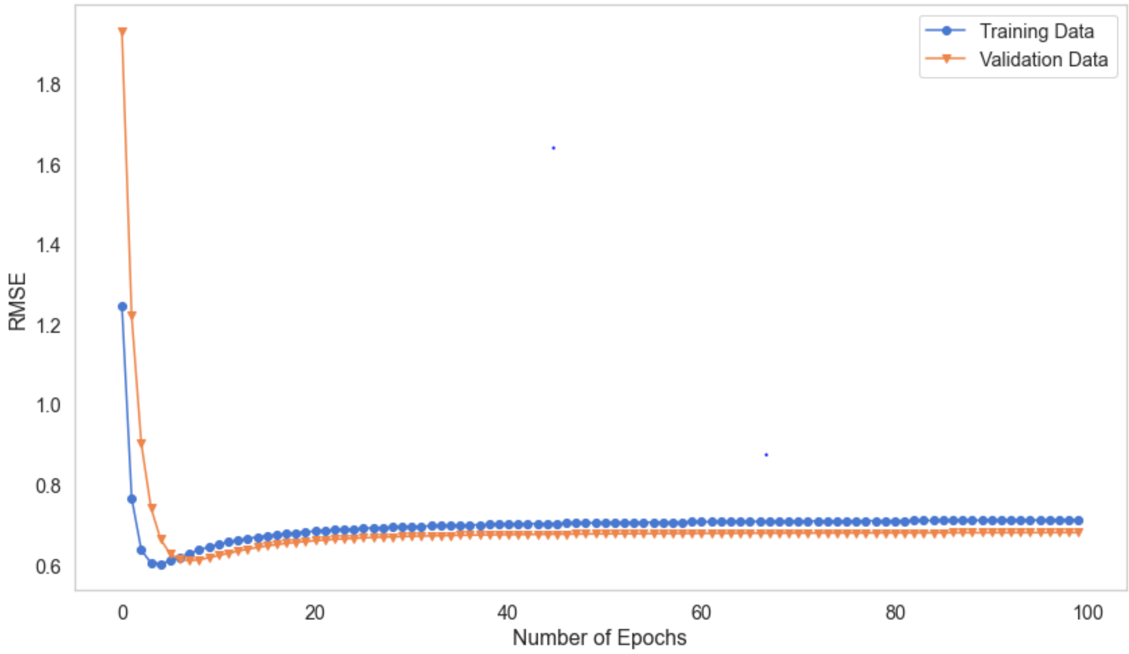
λ = 2

lambda: 2.0

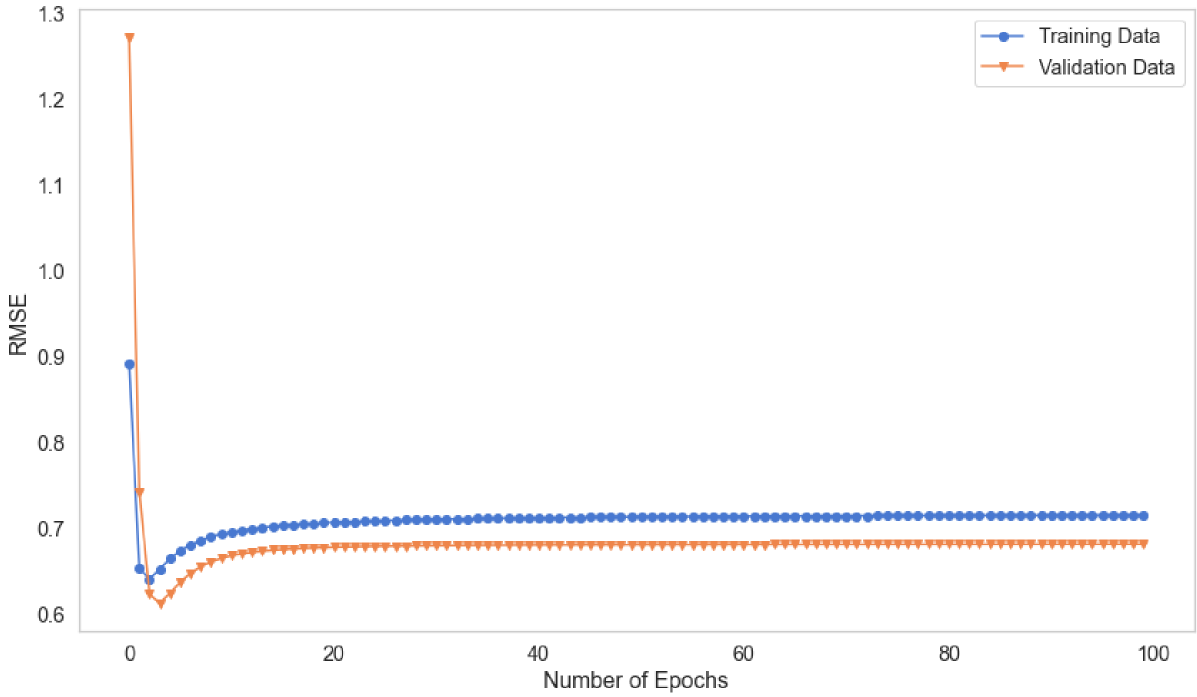


Shrey Shah
CWID - 20009523

λ = 5

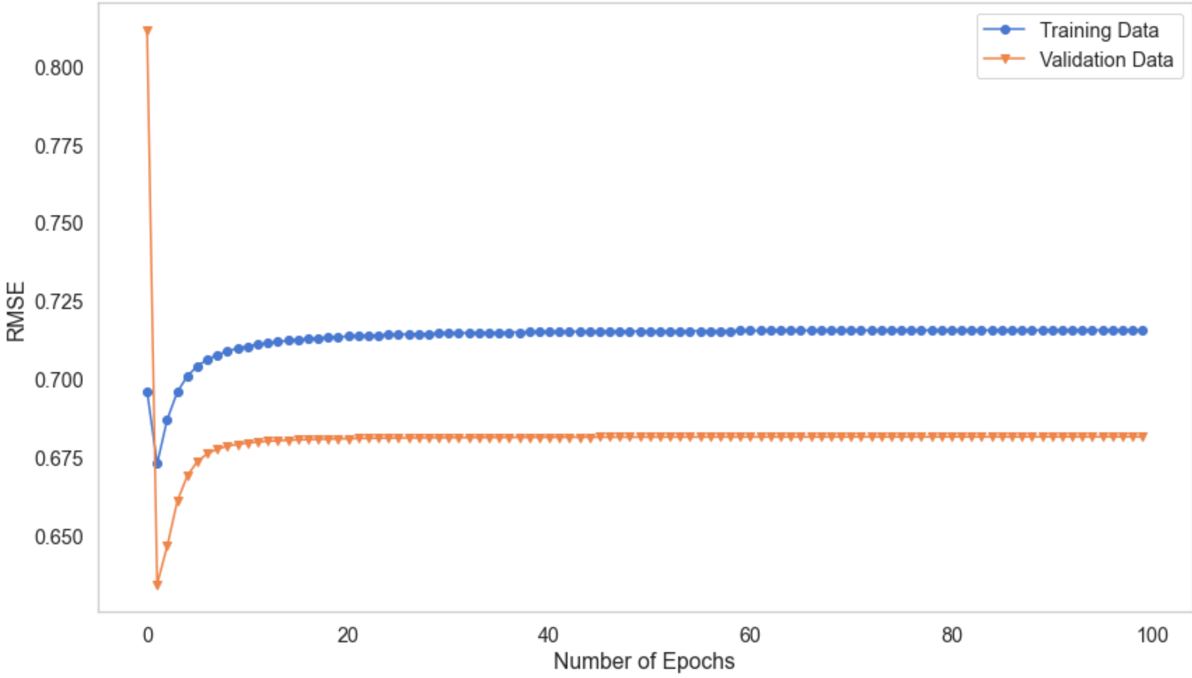


λ = 10



Shrey Shah
CWID - 20009523

λ = 20



λ = 50

Shrey Shah
CWID - 20009523