



THE UNIVERSITY OF
CHICAGO

Department of Statistics

MASTER'S THESIS PRESENTATION

WEIDI PAN

Department of Statistics
The University of Chicago

Collaborative Filtering and Music Personalization

WEDNESDAY, February 19, 2020, at 11:00 AM
Jones 304, 5747 S. Ellis Avenue

ABSTRACT

Collaborative Filtering (CF) remains one of the most important and popular approaches to building a recommender system, and it can be combined with content-based methods, such as NLP models, to form more complicated recommender systems. CF has two subcategories: memory-based algorithms and model-based algorithms. Memory-based CF algorithms can be further divided into user-based CF and item-based CF. In this report, I would like to explore both the memory-based and model-based approaches on the Million Song Dataset and outline the potential advantages and disadvantages of each method. My analysis suggests that, given the dataset of interest, the two memory-based methods are similar in terms of their predictive accuracy, and that the model-based method (Singular Value Decomposition) performs better than the memory-based methods both in terms of predictive accuracy and computational speed.