Parts Franck Report

Introduction

With the ever-increasing amount of online information, recommender systems have become a go-to strategy to overcome information overload. Recommender systems facilitate personalization through e-commerce and have grown significantly in terms of public awareness since its rise in the nineties, in parallel to the Web debut. Indeed, may conferences and workshops are exclusively dedicated to this topic, such as the *ACM Conference on Recommender Systems (RecSys)*.

Recommender systems are also very diverse because they may leverage various types of user-preferences and user-requirement data to make the recommendations. The most classical methods in recommender systems include collaborative filtering methods, knowledge-based methods, and content-based methods. They form the fundamental pillars of machine learning based recommender systems.

These methods evolved to fit various scenarios, data domains, and contexts, such as time, location (ie. gps signal) and social media information. Numerous breakthroughs have been found such as the application of Matrix Factorization (MF) and Restricted Boltzmann Machines (RBM), not only for movie recommendation during the 2006 Netflix contest, but also query log mining, news recommendations, and computational advertising.

Conventional methods such as MF are considered state of the art and have been favored over content-based methods, due to dealing better with sparsity, scalability, and predictions. However, they face several challenges such as the cold start problem (new items or users are difficult to predict), very high data sparsity (most of the items are not rated by a user), scalability issues with O(m\*n) complexity in MF, difficulty to boost diversity of the recommended items, and the obstacles to learn latent implicit features.

The last decade has witnessed the Deep Learning (DL) revolution with powerful applications in computer vision, speech recognition, and natural language processing. Currently, it is also revolutionizing the state-of-the-art recommendation architectures by bringing more opportunities to improve performance, which can be generically quantified through RMSE for rating predictions. Indeed, DL approaches have overcome some of the obstacles faced by conventional recommender systems. These approaches can capture the non-trivial and non-linear user-item relationships, and they allow for the representation learning of more complex abstractions as data representations in the higher layers. Further, DL identifies intricate relationships within the data from the abundant sources available at hand such as textual, contextual, and visual information. Thus, the number of published papers has increased exponentially in the recent years. Even RecSys started to organize frequent seminars and workshops on the matter since the year 2016. It is to be noted that DL recommenders also face the following challenges: lack of interpretability, large data requirements, extensive hyperparameter tuning.

As a result, we intend to compare the strengths and limitations of Matrix Factorization (MF), one of the state-of-the-art conventional recommender systems, with DL approaches such as AutoRec and GNN. Other DL algorithms have been studied but we limited our analysis to these two for brevity. First, we will undertake a survey of the literature and explain these methods. Then, we will briefly describe the datasets we used. Third we will describe our experiments with MF, Autorec, and GNC with the goal to optimize prediction rating in user-item context. Finally, we will discuss the results before touching upon the limitations.