Behavior's Study of some Classic SVD-Models with Noisy Data in Movie Recommender Systems

Cupertino Lucero-Álvarez[1,2], Perfecto Malaquías Quintero-Flores[2],
Edgar Moyotl-Hernández[3], Carlos Artemio Ortiz-Ramírez[1],
Patricia Mendoza-Crisóstomo[1], Maria Vázquez-Vázquez[1]

[1] Universidad Tecnológica de Izúcar de Matamoros,
Tecnologías de la Información,
Mexico

[2] Universidad Autónoma de Tlaxcala,
Facultad de Ciencias Básicas Ingeniería y Tecnología,
Mexico

[3] Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias Físico Matemáticas,
Mexico

{clucero, cortiz, pmendoza, mari.vazquez}@utim.edu.mx,
parfait.phd@gmail.com, moyotl@buap.edu.mx

Abstract. This paper presents a study about the behavior of three variants of the SVD algorithm in Collaborative Recommender Systems (CRS). For this, two MovieLens DataSets are used, and five variants in each DataSet with different degrees of randomness. Specifically, a comparison of the classic models is presented: Funk-SVD, Regularized-SVD, and Bias-SVD. The underlying idea is to observe that, as the degree of randomness in the data increases, the precision of the recommendations decreases, and the hidden relationships that may exist in the original data they get lost because of the noise. For this, we have configured two groups of experiments: in the first group, in each execution 10, 20 and 30 Latent Factors (LFs) were considered in the three models, while in the second group from 5 to 80 LFs were used in the regularized-SVD model. The prediction error was minimized using the MSE (Mean Square Error) metric and the ADAM optimizer. The results show that SVD with biases performs better, under the conditions of these experiments, and that noise affects the hidden relationships between the data.

Keywords: Collaborative recommendation systems, matrix factorization, singular value decomposition, latent factors, noisy data.

1 Introduction

Due to the vigorous growth of electronic commerce today, the need for efficient management of the Big Data generated is pressing. In the area of Recommendation Systems (RS), applications need efficient algorithms in the use of the computational