W inweb
Instituto Nacional de Ciência e Tecnologia para a Web

Department of Computer Science
Universidade Federal de Minas Gerais

CNPq
Conselho Nacional de Desenvolvimento
Científico e Tecnológico

Workshop on Recommender Systems
for a Massively Connected Society

TCC/TSI/TECC: Recommender Systems
Profs. Rodrygo Santos and Nivio Ziviani

FAPEMIG
Fundação de Amparo à Pesquisa do
Estado de Minas Gerais

Multi-Market Deal Size Prediction

Evelin C. F de Amorim <evelinamorim@ufmg.br>

Objective   We propose a method that predicts deal size considering that a deal can belong to more than one latent market.

Introduction

Daily-deals sites (DDSs) are popular web portals that offer discount coupons for services or products. Deal size prediction is a task that intent to predict the number of coupons solde in DDSs. We proposed a method for deal size prediction based on (Lacerda,2013), but we consider multi-market offers.

Related work

Lacerda (2013) proposed a strategy, called Competitive Business Market Prediction (CPMB), which is composed by three main steps: the first step divides the catalog of deals into markets by Latent Dirichlet Allocation (LDA); the second step performs regression training in each market; and the third step executes an compute the competition among the deals in the same market.

Proposed solution

Each deal presents textual features like: title, description, etc. After process textual features, Latent Dirichlet Allocation (LDA) method is applied in order to identified the markets. The weighting scheme for LDA is term spread(Lacerda,2013). After the processing of textual features, each deal is assigned to the most probable markets. Then the regression method is trained   in each market. Then a weighted mean is performed in order to find the deal sizes. Lacerda, also performed a step that consider the impact of markets in deal size. But we did no implement this step.

Experimental results

Dataset of 2590 instances crawled from Living Social: March 21st to July 3rd , 2011 (Byers et. al, 2012)

$$RRSE(y, \hat{y}) = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y_i} - y_i)^2}{\sum_{i=1}^{n}(\bar{y} - y_i)^2}}$$

Strategy	RRSE
Best-Market	1.03
All-Market	1.13
Thresh-Market (0.3)	1.06

Conclusions

Although, the error of our proposal is high than Lacerda, more tests must be performed. For instance, we can vary the threshold value.

Also we want to investigate whether errors are been introduced by markets that are not so probable.

Confidence interval will be computed as well, and we own another dataset from Groupon, which will enable us to perform further investigation.

References

Byers, J. W., Mitzenmacher, M., & Zervas, G. (2012,February). Daily deals: Prediction, social diffusion, and reputational ramifications. In *Proceedings of the fifth* ACM *international*
Lacerda, A.M. (2013). Revenue Optimization and Customer Targeting in Daily-Deals Sites. PhD. Thesis