Context-aware deal size prediction

Anisio Lacerda[1], Adriano Veloso[1], Rodrygo L. T. Santos[1], and Nivio Ziviani[1,2]

[1] Department of Computer Science
Universidade Federal de Minas Gerais
Belo Horizonte, MG, Brazil
{anisio,adrianov,rodrygo,nivio}@dcc.ufmg.br
[2] Zunnit Technologies
Belo Horizonte, MG, Brazil
nivio@zunnit.com

Abstract. Daily deals sites, such as Groupon and LivingSocial, attract millions of customers in the hunt for products and services at substantially reduced prices (i.e., deals). An important aspect for the profitability of these sites is the correct prediction of how many coupons will be sold for each deal in their catalog—a task commonly referred to as deal size prediction. Existing solutions for the deal size prediction problem focus on one deal at a time, neglecting the existence of similar deals in the catalog. In this paper, we propose to improve deal size prediction by taking into account the context in which a given deal is offered. In particular, we propose a topic modeling approach to identify markets with similar deals and an expectation-maximization approach to model intra-market competition while minimizing the prediction error. A systematic set of experiments shows that our approach offers gains in precision ranging from 8.18% to 17.67% when compared against existing solutions.

1 Introduction

In recent years, daily deals sites (or simply DDSs) such as Groupon[3] and LivingSocial[4] became an important group-buying alternative for both local merchants and consumers. While local merchants are mainly interested in disseminating their brand to increase revenue, potential consumers seek discounted prices for products and services as diverse as restaurant meals, theater tickets, etc. In this business model, the DDS operates as a mediator for local merchants (sellers) to negotiate a deal (product or service) that is sold to consumers (buyers). In this case, the profitability of the DDS depends directly on two factors: (i) the commission associated with a deal, and (ii) the number of coupons that are actually sold for the deal, a quantity commonly referred to as the deal size. Different from commissions, which are governed by business decisions, the task of predicting the deal size can be modeled as a machine learning regression problem [8].

While being of paramount importance for the success of DDSs, accurately predicting deal sizes is surrounded by challenges. First, the catalog of deals is

[3] http://www.groupon.com
[4] http://www.livingsocial.com

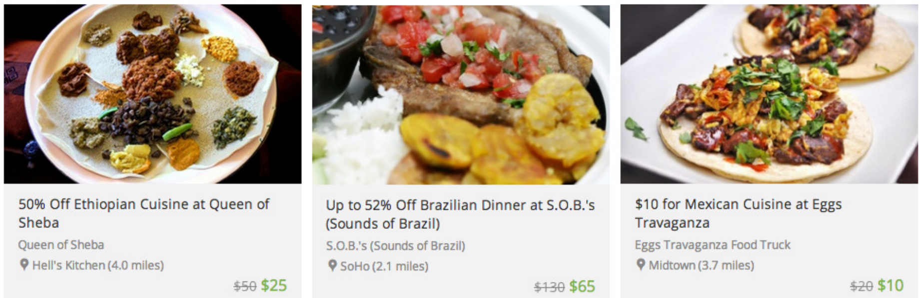

Fig. 1: The deal size depends on the deal and the presented alternatives.

usually available for only a limited time frame, which varies from 4 to 5 days on average [8]. This may compromise the amount of historical data that is available for learning predictors, harming the effectiveness of algorithms such as support vector regression (SVR [3,11]). Second, deals may compete among themselves for consumer preference. For instance, consider the example in Fig. 1, which illustrates a commonly observed case in which a consumer is looking for discounts in restaurants and has a limited budget of $50. In this case, she would arguably prefer an Italian dinner over a Kebab, or vice-versa, but she is unlikely to increase her budget in order to buy both deals. An even worse outcome may happen when the similarity between the two services is so high that causes hesitation, leading the consumer to abandon the DDS without buying any of the competing deals.

Existing solutions to deal size prediction [8,14] produce global predictors that often neglect complex interactions between deals, such as competition for customer preference. In contrast, we propose to model the attractiveness of a deal relatively to other available deals from the same market. In particular, we propose a topic modeling approach to identify sets of deals that are likely to attract the interest of similar consumers. Furthermore, we propose a context-aware expectation-maximization approach to deal size prediction by considering (i) features associated with the target deal, and (ii) contextual features associated with the other deals in the same market of the target deal. To the best of our knowledge, our proposed approach is the first learning model that takes into account the whole catalog of deals when performing deal size prediction.

To assess the effectiveness of our proposed approach, we perform a systematic evaluation involving real usage data obtained from major DDSs such as Groupon and LivingSocial. In order to evaluate the extent to which our market segmentation strategy is language-dependent, we perform additional experiments with usage data obtained from Peixe Urbano,[5] the largest Brazilian DDS. The results attest the effectiveness of our proposed approach, with precision improvements ranging from 8.18% to 17.67% when compared to existing deal size predictors.

[5] http://www.peixeurbano.com.br

2 Related work

DDSs have recently attracted the attention of researchers in multidisciplinary fields. Regarding the economic aspect of DDSs, Byers et al. [7] presented evidence that Groupon strategically optimizes their deal offerings, giving customers incentives other than price to make a purchase, including deal scheduling and duration, deal featuring, and limited inventory. Groupon's business model was further examined by Arabshahi [2]. An empirical analysis of the experience of merchants that used Groupon was performed by Dholakia [10]. Several studies addressed the propagation effect of daily deals in online social networks. More specifically, they were interested in assessing the impact that a deal had on a merchant's subsequent ratings in social review sites such as Yelp. Different from past research that observed a decrease in the ratings for merchants using Groupon [9], Potamias [17] argued that this effect was overestimated. Finally, Kumar and Rajan [13] analyzed, among other economic aspects, the profitability of social coupons and concluded that they yield profits for merchants.

Another line of research addressed algorithmic problems in DDSs. In particular, these data-driven approaches focus on identifying and understanding the main characteristics of DDSs by analyzing historical purchase data. Two main interrelated problems arise in this scenario: (i) deal ordering, aimed at selecting a set of deals that should be featured in the DDS catalog on a given day, and (ii) deal size prediction, aimed at estimating the number of coupons that are expected to sell for a given deal. As previously discussed, the deal size prediction problem is the focus of this paper. In this line of research, Ye et al. [22] modeled the popularity of group deals as a function of time. Byers et al. [8] modeled deal size prediction as a linear combination of deal features and used ordinary least squares regression to fit their model. Instead of learning a single, global predictor, Lappas and Terzi [14] proposed to learn multiple predictors, one for each market identified using a hierarchical clustering algorithm.

Similarly to the prediction approach of Lappas and Terzi [14], we also seek to identify multiple markets and learn market-targeted deal size predictors. However, in contrast to their approach, we leverage several weighting schemes and structural properties of deals as discriminative features for market identification. Moreover, we introduce a normalization factor learned via expectation-maximization to account for competing interactions between deals from the same market in order to produce context-aware predictions. In our investigations in Section 5, both the global prediction approach of Byers et al. [8] and the segmented prediction approach of Lappas and Terzi [14] are used as baselines.

3 Context-aware deal size prediction

In this section, we introduce our context-aware deal size prediction approach. Firstly, we present a topic identification strategy to group deals into markets, which determine the context for each available deal. Then, we present a deal size prediction strategy that employs multiple SVR predictors: there are as many predictors as markets, and each predictor is specifically designed to one market. Finally, we present a contextual expectation-maximization strategy that reduces the prediction error by taking into account the competition between deals in the same market and also the representativeness of different markets.

3.1 Identifying markets

Discovering meaningful markets from textual features associated with the deals is an important step for determining the context of competition among deals. In the following, we detail our approach for representing deals as well as for automatically identifying markets via latent topic modeling.

Representing deals. Our approach represents each deal d as a vector $\boldsymbol{d}_s$ in an n-dimensional space $\{t_1, t_2, \ldots, t_n\}$, where n is the number of unique terms in a given feature space s. In particular, we consider four different features spaces, comprising terms appearing on the *merchant's name*, the *title of the deal*, the *description of the deal*, or *all of these fields* concatenated in a single space.

In order to weigh the relative importance of each term t in a given deal vector $\boldsymbol{d}_s$, we consider four alternative weighting schemes: the *term frequency* TF, denoting the raw frequency of t in d; the *term spread* TS, denoting the spread of t in d, as a measure of the descriptive power of t; and the products TF$\times$IFF and TS$\times$IFF, with the *inverse feature frequency* IFF denoting the rarity of t among all cataloged deals represented in the feature space s.

Identifying latent markets. Our proposed approach defines a market as a set of deals that are likely to attract the interest of a similar group of customers. Under the assumption that customers have a limited budget, our intuition is that deals belonging to the same market are more likely to compete for customer preference. A simple strategy to market identification would be to analyze the customers' purchase history, in order to identify deals that were purchased by a similar set of customers. Unfortunately, as discussed in Section 1, such historical purchase data is very limited due to the scarceness of recurrent or regular customers. As an alternative, we propose a content-based approach to market identification using latent Dirichlet allocation (LDA [5]). LDA is a generative model used to identify latent topics in textual documents, and has been largely used in a variety of tasks, including matrix factorization [1], influential user identification [21], tag recommendation [12], and word sense disambiguation [6]. In our particular case, using LDA to identify latent markets overcomes the lack of historical purchase data by instead leveraging the textual representation of each of the available deals. As an illustrative example, Table 1 shows markets identified from one of the DDS datasets used in our investigations in Section 5.

3.2 Predicting deal size

In order to predict the deal size, we learn multiple predictors, each one targeted to a different market, identified according to the approach described in Section 3.1.

Table 1: Top five terms from each of $k = 5$ markets in the Groupon dataset.

"Gym"	"Hair Salon"	"Sports"	"Dentistry"	"Ice Cream"
class	salon	camping	dental	tour
fit	hair	week	teeth	chocolate
body	look	sport	care	cake
train	services	day	value	sweet
workout	cut	academia	whitening	ice

To account for the competition within each market, these initial predictions are further adjusted through an iterative expectation-maximization procedure.

Specialized SVR. Support vector regression (SVR [11]) is an established non-linear regression technique that has been applied successfully to a variety of numeric prediction problems. In order to apply SVR to deal size prediction, we represent deals as follows. Let $\mathcal{D} = \{\mathcal{D}_{m_1}, \mathcal{D}_{m_2}, \ldots, \mathcal{D}_{m_k}\}$ be the collection of all past deals, and each $\mathcal{D}_{m_i} = \{d_1, d_2, \ldots, d_q\}$ be a partition of $\mathcal{D}$ composed of all q deals that belong to market m_i. Further, each deal d_j is represented as a feature vector, using features generally available from DDSs, such as the deal face and discounted values, the day of the week when the deal is launched, etc.

The training set used to build an SVR predictor to market m_i is composed of (deal, size) pairs of the form $\{(d_1, s_1), (d_2, s_2), \ldots, (d_q, s_q)\}$, that is, each pair is composed of a deal $d_j \in \mathcal{D}_{m_i}$ and its corresponding size s_j. For each deal $d_p \in \mathcal{D}_{m_i}$, a specialized SVR predictor takes the form $f(d_p) = \langle w, \Phi(d_p)\rangle + b$, where $w \subset \Re^n$, $b \subset \Re$, and Φ denotes a nonlinear transformation from $\Re^n$ to a high-dimensional space. The SVR objective is to find the minimum value of w and b by solving the following regularized optimization problem:

$$\min_w \frac{1}{2} w^T w + C \sum \xi_\epsilon(w; d_j, s_j), \tag{1}$$

$$\xi_\epsilon(w; d_j, s_j) = \max(|w^T d_j - s_j| - \epsilon, 0)^2, \tag{2}$$

where $C > 0$ is the regularization parameter, and ξ_ϵ is the ϵ-insensitive loss associated with (d_j, s_j), with ϵ given so that the loss is zero when $|w^T d_j - s_j| \geq \epsilon$ [19]. We use a radial basis function (RBF) as the transformation function Φ, and set all other parameters through cross-validation [15].

The CBMP model. The specialized SVR predictor does not take into account complex interactions that may exist between deals in the catalog. For instance, deals in the catalog of the day, denoted S, may compete with each other. As a result, SVR predictions may be overestimated. In order to model competition, we first partition the catalog into k markets, so that $S = \{S_{m_1}, S_{m_2}, \ldots, S_{m_k}\}$. Every deal in S must be assigned to one of these markets, and we say that S_{d_j} is the market to which deal d_j is assigned. Given the SVR prediction $f(q)$ for any deal $q \in S$, the estimated size of deal $d_j \in S$ is given as a combination of two factors: the SVR prediction for d_j, and an average factor that encompasses all deals within the same market. In order to combine these two factors, we introduce the following expectation-maximization (EM) formulation:

$$E: \quad \rho(S) = \frac{\sum_{q \in S} \sigma(q)}{\sum_{d_j \in C} \sigma(d_j)}, \tag{3}$$

$$M: \quad \sigma(d_j) = \alpha \times \frac{\sum_{q \in S_{d_j}} f(q) \times \rho(S_{d_j})}{|S_{d_j}|} + (1 - \alpha) \times f(d_j), \tag{4}$$

where $\alpha = 0.5$ to weigh equally both factors, and $\rho(S_{d_j})$ is an unknown parameter which re-scales the representativeness of market S_{d_j}. This EM procedure uses the training set to find the value for $\rho(S_{d_j})$ that minimizes the loss function:

$$eqn : rmse(y, \hat{y}) = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}, \tag{5}$$

where n is the number of predictions performed, and y_i and $\hat{y}_i$ are the actual and the predicted deal sizes, respectively. Intuitively, this prediction model, named Competitive Business Market Predictor (CBMP), accounts for intra-market competition by re-scaling the predicted size of each deal relatively to the predicted size of other deals from the same market. In the next sections, we assess the effectiveness of the CBMP model compared to both a global as well as market-targeted predictors which do not account for competing relationships.

4 Experimental setup

This section details the experimental setup that supports the validation of our proposed context-aware deal size prediction approach, including the datasets, the baselines, and the evaluation procedure that we use.

Daily deals datasets. Our evaluation uses data collected from three commercial DDSs: Groupon, LivingSocial, and Peixe Urbano. Groupon and LivingSocial were crawled using as seeds the datasets used by Byers et al. [8]. In particular, Groupon was crawled from Jan 3rd, 2011 to Jul 3rd, 2011, while LivingSocial was crawled from Mar 21st, 2011 to Jul 3rd, 2011. In both cases, the collected data includes English textual features used to group deals into markets, as described in Section 3.1. For Groupon, we collected a total of 16,409 deals comprising 119,525 unique terms. For LivingSocial, we obtained 2,610 deals with 19,102 distinct terms. In addition, to enable our basic prediction model using SVR, we collected a number of non-textual deal features, including (1) the price of the deal, (2) the price after the discount, (3) the tipping point of the deal (i.e., minimum number of coupons that must be sold to enable the deal), (4) the day of the week in which the deal was launched, (5) the category of the deal, (6) the city in which the corresponding merchant is located, and boolean values indicating (7) whether the deal is running for multiple days, (8) whether the deal is featured on the DDS website, and (9) whether the inventory is limited. Finally, for Peixe Urbano, the same features were obtained directly through an API made available to us. Textual features in Portuguese were obtained for 4,309 deals during the entire year of 2012, comprising a total of 31,163 unique terms.

Prediction baselines. We compare the effectiveness of CBMP, our context-aware deal size prediction model, against the following baseline predictors:

– Global Predictor (GLPR) [8], learned using the whole training set, ignoring the existence of markets. The size of an arbitrary deal q is given as:

$$\sigma(q) = 2^{\beta_0 + \sum \beta_i \times f_i}, \tag{6}$$

where features f_i correspond to the ones described earlier in this section and weights β_i are determined using ordinary least squares [15].
– One Predictor per Business Market (OPBM) [14], learned using an SVM-based regression method proposed by Shevade et al. [18]. In particular, they introduced a two-layer clustering algorithm that is based on LDA [5] and a flat clustering algorithm based on the Kullback-Liebler distance to separate deals into markets [16]. Note that, in contrast to our proposed model, OPBM does not exploit competing relationships between deals in a given market.

Evaluation procedure. Our evaluation uses the interleaved test-then-train methodology [4], in which each deal in the catalog on day t is evaluated, and then is included into the historical data that becomes available at day $t+1$. Our intention is to mimic as close as possible the production system of a DDS, which predicts the size of each available deal in the beginning of the day, and incorporates the feedback received on these deals in the end of the day, to be used in the predictions of the next day. We evaluate the effectiveness of all predictors using the root mean squared error (RMSE) [15]. As shown in Equation (5), $\hat{y}_i$ is the predicted value of the i-th sample, while y_i is the corresponding true value. Finally, n is the number of instances in the dataset for which the prediction is performed. The results to be reported correspond to an average over all days. Statistical significance was verified using a paired t-test with $p < 0.05$.

5 Experimental results

In this section, we empirically validate our proposed context-aware deal size prediction model in contrast to existing deal size predictors from the literature. In particular, we aim to answer the following research questions:

Q1. How effective is CBMP compared to existing deal size predictors?
Q2. How effective is our topic modeling approach to market identification?
Q3. How effective are our various strategies for deal representation?

Table 2: RMSE figures for CBMP and the GLPR and OPBM baselines.

	GLPR	OPBM	CBMP$_l$	Δ_{GLPR}	Δ_{OPBM}	CBMP$_c$	Δ_{GLPR}	Δ_{OPBM}
Groupon	1.3864	1.3544	1.1332	17.7%	16.3%	1.2563	8.7%	7.2%
LivingSocial	1.1287	1.1112	1.0203	9.6%	8.2%	1.0931	3.2%	1.6%
Peixe Urbano	1.2956	1.2575	1.1430	11.8%	9.1%	1.1841	8.6%	5.8%

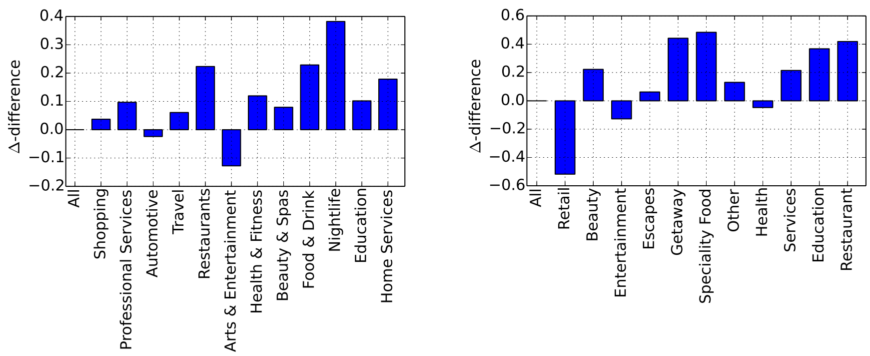

Fig. 2: Market-specific prediction model vs. global model.

To answer question Q1 and Q2, we contrast the effectiveness of our CBMP model to GLPR as a global, market-agnostic predictor, as well as to OPBM as a local, market-aware predictor that does not exploit the competing nature of deals in the context of an individual market. Table 2 shows the prediction performance of these three models across the three considered DDS datasets (Groupon, LivingSocial, and Peixe Urbano) in terms of RMSE. Our CBMP model is deployed using either latent topics (CBMP$_l$) or explicit categories (CBMP$_c$) for market identification. The CBMP$_l$ variant identifies latent markets (50 for Groupon; 30 for LivingSocial and Peixe Urbano) with deals represented by a concatenation of all their terms weighted using TF, which was the best performing setting identified during training.[6] For each of the two variants of CBMP, namely, CBMP$_l$ and CBMP$_c$, the additional columns Δ_{GLPR} and Δ_{OPBM} show the percentage RMSE improvement compared to GLPR and OPBM, respectively.

From Table 2, we first observe that both variants of our CBMP model outperform both deal size prediction baselines. In particular, CBMP$_c$ outperforms the global GLPR predictor with significant improvements in RMSE ranging from 3.2% (LivingSocial) to 8.7% (Groupon). These gains are further illustrated by the breakdown analysis shown in Fig. 2 for deals across multiple categories of Groupon and LivingSocial.[7] In the figure, a positive delta indicates an improvement of CBMP against GLPR, whereas a negative delta indicates otherwise.

[6] A complete analysis of the impact of different deal representations for a varying number of target markets is presented later in this section.

[7] Results on Peixe Urbano show similar trends and are omitted for brevity.

From the figure, we observe that most of the observed differences are positive, indicating that it is indeed better to use multiple market-specific predictors than a single, global predictor for all markets. Notable exceptions can be observed for categories such as "Retail" and "Entertainment", when a global predictor performs better. Such categories are arguably less cohesive, which may indicate weaker intra-market competition relationships. Further exploiting such a nuanced view of markets is a direction for future investigation.

A promising alternative to arbitrarily defined markets is the latent topic identification approach employed by CBMP_l. Indeed, as shown in Table 2, CBMP_l further improves compared to CBMP_c, which is based on explicit categories. In particular, CBMP_l significantly outperforms the global GLPR predictor by up to 17.7% (Groupon). Compared to the OPBM baseline, which also leverages topic modeling to produce market-specific deal size predictions, our approaches are also effective, with gains ranging from 8.2% (LivingSocial) to 16.3% (Groupon) for CBMP_l and 1.6% (LivingSocial) to 7.2% (Groupon) for CBMP_l. Recalling question Q1, these results attest the effectiveness of our proposed context-aware deal size prediction model compared to both global as well as market-specific, content-agnostic predictors. Moreover, recalling question Q2, the comparison between the two variants of our model also attest the effectiveness of our topic modeling approach to identify latent markets, which outperforms a hard partition of markets based on the category of each deal in the catalog.

To address question Q3, we further evaluate our topic modeling variant (henceforth referred to as CBMP) in light of different strategies for deal representation. To this end, we break down the performance of CBMP for different feature spaces (merchant's name, deal title, deal description, concatenation of all terms), weighting schemes (TF, TS, TF$\times$IFF, and TS$\times$IFF), and target number of markets. In particular, Figs. 3, 4, and 5 show the results of this breakdown analysis for the Groupon, LivingSocial, and Peixe Urbano datasets, respectively.

Regarding the impact of different feature spaces, we observe a generally superior performance of CBMP in denser spaces, such as those induced with terms obtained from the description or the concatenation of all terms comprised by each deal. In addition, of these two feature spaces, the concatenation of all terms appears to be less sensitive to the variation in the number of chosen markets. Considering the other two spaces, merchant's name and deal title, both yield generally similar prediction performances across the three datasets. Regarding the different weighting schemes considered, TF and TS are consistently the best performers across all datasets. Furthermore, we can also see that the weighting schemes that are combined with IFF are more sensitive to the number of markets. For instance, when considering the deal description on Groupon (Fig. 3), or the deal title on LivingSocial (Fig. 4) and Peixe Urbano (Fig. 5).

Recalling question Q3, the results in Figs. 3, 4, and 5 show that denser feature spaces with pure frequency-based weighting functions are generally preferred for representing deals within our latent market identification approach. The target number of markets, on the other hand, is a key parameter that must be carefully tuned when deploying our proposed context-aware deal size prediction model for different datasets. Finally, it is also worth noting that our approach performs effectively for datasets in different languages, as exemplified by the English and Portuguese DDS datasets considered in our investigations.

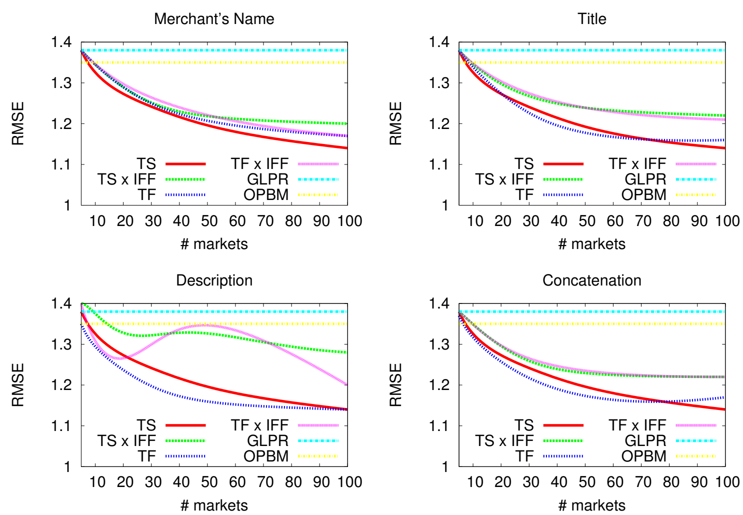

Fig. 3: RMSE on Groupon with varying number of markets.

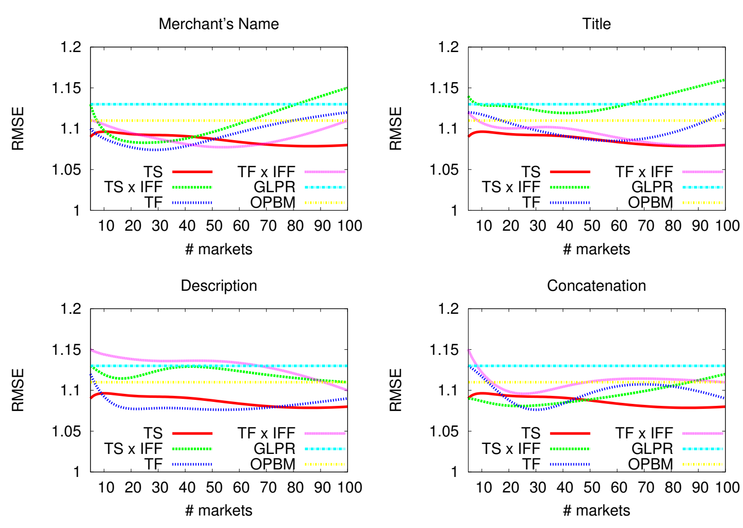

Fig. 4: RMSE on LivingSocial with varying number of markets.

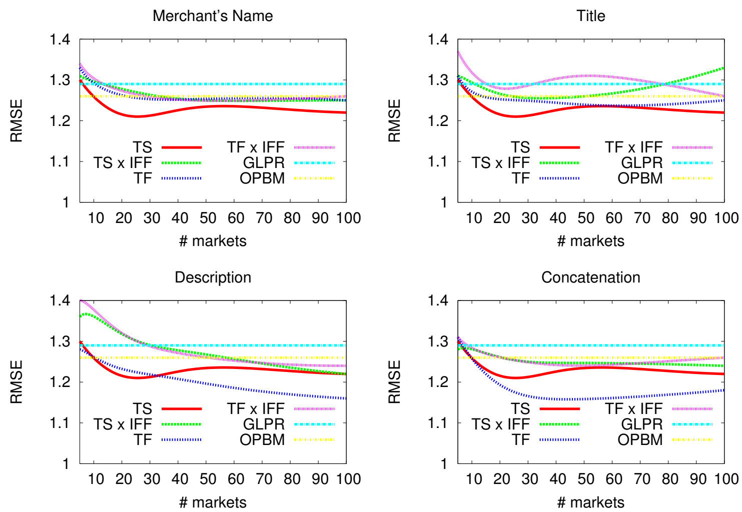

Fig. 5: RMSE on Peixe Urbano with varying number of markets.

6 Conclusions and future work

We introduced CBMP, a novel context-aware deal size prediction model. Predicting the size of a deal (i.e., the number of coupons that will be sold for the deal) is a crucial task for the profitability of DDSs. Our proposed model improves upon previous approaches by exploiting competition relationships that may arise among deals in the same target market (e.g., two restaurant deals). In particular, we proposed a topic modeling approach to identify latent markets based solely on the textual content of deals. Besides identifying more cohesive markets, this content-based approach is particularly suitable for the dynamic nature of DDSs, where the volatility of the available deals precludes an effective use of historical purchase data. Based upon the identified markets, we proposed an expectation-maximization formulation to re-scale the predicted size of a deal in light of the predicted size of other competing deals from the same market. Experiments on three large-scale datasets collected from commercial DDSs attested the effectiveness of our proposed approach, with significant gains in prediction accuracy ranging from 8.2% to 17.7% over previously proposed approaches.

We exploited the relationship among deals by using the concept of markets, which is useful to model flips in consumer behavior. However, such behavior presents many other aspects that have been studied in behavioral economics and may be used in the context of DDSs for deal size prediction. For instance, anchoring [20] states that the first product shown influences the buying decisions for subsequently shown products, which are compared to the first one. Hence, in the context of DDSs, the ordering of the products presented in the web page may affect deal size, i.e., the top product may also have an anchoring effect.

Acknowledgements

We thank the partial support given by the Brazilian National Institute of Science and Technology for the Web (grant MCT-CNPq 573871/2008-6) and authors' individual grants and scholarships from CNPq and CAPES.

References

1. D. Agarwal and B.-C. Chen. fLDA: matrix factorization through latent dirichlet allocation. In *ACM WSDM*, pages 91–100, 2010.
2. A. Arabshahi. Undressing groupon: An analysis of the groupon business model. 2011. http://www.ahmadalia.com/blog/2011/01/undressing-groupon.html.
3. D. Basak, S. Pal, and D. C. Patranabis. Support vector regression. *Neural Information Processing-Letters and Reviews*, 11(10):203–224, 2007.
4. A. Bifet, G. Holmes, R. Kirkby, and B. Pfahringer. MOA: Massive online analysis. *The Journal of Machine Learning Research*, 11:1601–1604, 2010.
5. D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.
6. J. L. Boyd-Graber, D. M. Blei, and X. Zhu. A topic model for word sense disambiguation. In *ACM ACL*, pages 1138–1147, 2010.
7. J. Byers, M. Mitzenmacher, M. Potamias, and G. Zervas. A month in the life of groupon. *CoRR*, abs/1105.0903, 2011.
8. J. Byers, M. Mitzenmacher, and G. Zervas. Daily deals: prediction, social diffusion, and reputational ramifications. In *ACM WSDM*, pages 543–552, 2012.
9. J. W. Byers, M. Mitzenmacher, and G. Zervas. The groupon effect on yelp ratings: A root cause analysis. In *ACM EC*, pages 248–265, 2012.
10. U. M. Dholakia. How effective are groupon promotions for business. 2010. http://www.ruf.rice.edu/~dholakia.
11. H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik. Support vector regression machines. *NIPS*, pages 155–161, 1997.
12. R. Krestel, P. Fankhauser, and W. Nejdl. Latent dirichlet allocation for tag recommendation. In *ACM RecSys*, pages 61–68, 2009.
13. V. Kumar and B. Rajan. Social coupons as a marketing strategy: A multifaceted perspective. *Journal of the Academy of Marketing Science*, 40(1):120–136, 2012.
14. T. Lappas and E. Terzi. Daily-deal selection for revenue maximization. In *ACM CIKM*, pages 565–574, 2012.
15. T. M. Mitchell. Machine learning. 1997. *Burr Ridge, IL: McGraw Hill*, 45, 1997.
16. D. Pinto, J.-M. Benedí, and P. Rosso. Clustering narrow-domain short texts by using the kullback-leibler distance. In *CICLing*, pages 611–622, 2007.
17. M. Potamias. The warm-start bias of yelp ratings. *CoRR*, 2012.
18. S. K. Shevade, S. S. Keerthi, C. Bhattacharyya, and K. R. K. Murthy. Improvements to the SMO algorithm for svm regression. *Neural Networks, IEEE Transactions on*, 11(5):1188–1193, 2000.
19. A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, 2004.
20. A. Tversky and I. Simonson. Context-dependent preferences. *Management science*, 39(10):1179–1189, 1993.
21. J. Weng, E.-P. Lim, J. Jiang, and Q. He. Twitterrank: finding topic-sensitive influential twitterers. In *ACM WSDM*, pages 261–270, 2010.
22. M. Ye, T. Sandholm, C. Wang, C. Aperjis, and B. A. Huberman. Collective attention and the dynamics of group deals. In *WWW*, pages 1205–1212, 2012.