Towards Context Aware Food Sales Prediction

Indrė Žliobaitė, Jorn Bakker, Mykola Pechenizkiy

Eindhoven University of Technology, Department of Computer Science
P.O. Box 513, 5600 MB Eindhoven, NL

Abstract

Sales prediction is a complex task because of large number of factors affecting the demand. We present a context aware sales prediction approach, which selects the base predictor depending on the structural properties of the historical sales. A case study of a food wholesaler indicates that moving average prediction can be outperformed by intelligent methods, if a proper categorization is in place.[1]

1 Introduction

Demand prediction is essential part of business planning. An accurate and timely sales prediction is essential for stock management. That is a crucial part for food wholesales and retail profitability. The stock includes large assortment of goods, some of them require special storage conditions, some are quickly perishable.

There are general and product specific causes of the demand fluctuation. The variations in consumer demand may be influenced by price (change), promotions, changing (rapid or gradual, global or local) consumer preferences or weather changes. Furthermore, a large share of the products is sensitive to some form of a seasonal change. Seasonal changes occur due to different cultural habits, national, school or religious holidays, fasting. All these factors imply that some types of products have high sales during a limited period of time.

Although seasonal patterns are expected, the predictive features that define these seasons are not always directly observed. Therefore, fluctuations in sales which are accommodated by the changing seasons are often difficult to predict. Besides, the historical data is often highly imbalanced, i.e. occasion specific products would have only a few weeks of the sales peaks per year.

The main idea of the context aware approach that we propose in this work is to select the predictor based on the structural properties of time series. Different products have different sales behavior and different dependence on calendar events (seasonality). If we can identify and extract distinct categories of products, specific input data construction procedures and specific predictors could be employed for each category. One could argue, that an ensemble approach with a rigorous feature selection and predictor selection [2] does that automatically. However, this approach has limitations with respect to a given food sales prediction problem. First of all, the data is noisy and relatively short. For example, a particular food wholesaler company keeps only two years record in their transactional database. If a product is seasonal and peaks only once per year for a particular event, we would have only one or two positive examples in the historical data. By defining the context, we filter out a part of noise. Secondly, some series share common patterns. For example, New Year peaks are common for a large subset of products. By categorizing the time series based on their structural properties, we narrow down the job for the particular predictor, allowing to focus on the peculiarities of a particular series.

2 Our approach

Fig. 1 presents a snapshot (one time step) of the context aware prediction approach (CAPA) we developed.

CAPA consists of two blocks - training (offline) and operational (online). Here we highlight only the operational online part and refer an interested reader to the full paper [1] for the detailed description of how

[1] This is an extended abstract of the full paper accepted to the Workshop on Domain Driven Data Mining, at IEEE ICDM 2009 [1]).

the model is trained and parameterized offline. Assume that the model has already been trained offline, i.e. the categories have been fixed, mapping of time series to the categories is established, "local" expertise of each predictor is known.

Let us take a particular product we are interested to predict online. First, we extract structural features from the original sales time series of the product. Then, we assign the product to one of the categories. We pick a base predictor (a regressor learner) specific to a particular category and select input feature sets, relevant to a particular category (that could be historical sales, calendar events, weather, and promotions). That is the context aware part of the approach. The contexts are specified by predefined categories. Having the original series, the base learner and the input features we can cast the prediction.[2] The prediction output now can be used for a decision making in a business process.

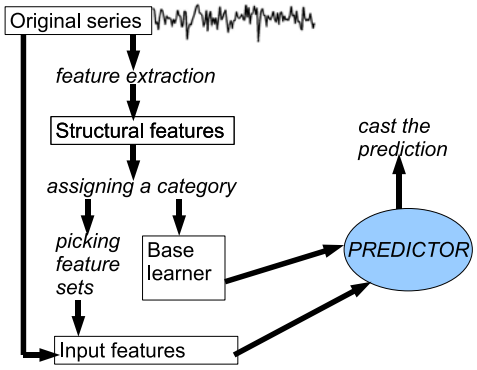

Figure 1: Online operation of CAPA

3 Results with a case study and conclusions

We conducted a case study of Sligro, the company engaged into food wholesales. In their business process the predictions are made using moving average but human corrections are made on top of the obtained predictions. Human factor is a volatile part of the decision making. Thus we aimed to summarize the seasonality related factors to build intelligent models with integrated context knowledge.

We performed a set of controlled experiments to justify our claims regarding context aware prediction. Our experimental field consists of 538 product sales quantities (119 weekly aggregates over all outlets) over two years period. Experimental scenario consisted of three parts: selection the base predictors, learning categorization rules and testing final model accuracies. In the first part we selected the base predictors and showed that there exist product subsets on which it is *possible* to outperform baseline predictor. In the second part we aimed to learn the dependencies between product categorization accuracies and sales prediction accuracies applying two different approaches. In the third part we tested the final model accuracies.

Our results demonstrated that distinct categories exist, where the intelligent learners (a regression equipped with feature selection mechanism in our case) can outperform naive predictors if online categorization is accurate enough. However, currently employed time-series categorization approach was not accurate enough. Further improvement of the obtained categorization accuracy could be achieved by finding more representative structural features, introducing multiple category assignments, adding more domain knowledge to the selection of the base predictors.

It would be also interesting to analyze the semantic relations between the products and product groups in a context of sales prediction. Hierarchical prediction would give sales estimates on aggregated level, product groups and individual products simultaneously.

CAPA is generic in a sense that we can reassign the product to another category at each time step based on recent development of sales. For example, a new marketing strategy cames into play promoting a particular beer to be the beer of rainy days. Therefore, next practical step would be to employ the approach in a dynamic setting.

References

[1] I. Žliobaitė, J. Bakker, and M. Pechenizkiy, "Towards context aware food sales prediction," in *DDDM'09, Workshops Proceedings of the 9th IEEE International Conference on Data Mining (ICDM 2009)*, IEEE Computer Society, 2009.

[2] P. Meulstee and M. Pechenizkiy, "Food sales prediction: "If only it knew what we know"," in *DDDM'08, Workshops Proceedings of the 8th IEEE International Conference on Data Mining (ICDM 2008)*, pp. 134–143, IEEE Computer Society, 2008.

[2] Note, that the predictor does not use the structural features as were used for assigning the product to a category.