Multi-View LS-SVM for Temperature Prediction

Lynn Houthuys, Zahra Karevan, and Johan A. K. Suykens
ESAT-STADIUS, KU Leuven. Kasteelpark Arenberg 10 B-3001 Leuven, Belgium

Abstract. In multi-view regression, the input data can be represented in multiple ways or views. The aim is to increase the performance of using only one view by taking into account the information available from all views. We introduce a novel multi-view regression model called Multi-View Least Squares Support Vector Machines (MV LS-SVM). This work was motivated by the challenge of predicting temperature in weather forecasting. Black-box weather forecasting deals with a large number of observations and features and is one of the most challenging learning tasks around. In order to predict the temperature in a city, the historical data from that city as well as from the neighboring cities are taken into account. We use MV LS-SVM to do temperature prediction by regarding each city as a different view. Experimental results on the min. and max. temperature prediction in Brussels, show the improvement of the multi-view method with regard to previous work and that it is competitive to the existing state-of-the-art methods in weather prediction.

1 Introduction

The multi-view [1] method proposed in this paper is called *Multi-View Least Squares Support Vector Machines (MV LS-SVM) Regression.* It is cast in the primal-dual setting typical to Least Squares Support Vector Machines [2] where the separate models for each view are combined in the primal objective function so that information from other views is taken into account during training.

We focus on MV LS-SVM regression for black-box weather forecasting. Accurate weather prediction is a challenging problem that can influence our daily lives in different ways. Lazo et al. reported that the U.S. public obtains more than 300 billion forecasts with a total estimated value of $31.5 billion each year [3]. In order to forecast a weather condition in a particular city, the historical data of some nearby cities have been taken into account. Instead of simply concatenating the feature vectors of all cities, this paper regards each city as a separate view. By using the novel MV LS-SVM regression method, the temperature influences of each city can be modeled separately, with different parameters for each city, while the coupling term enforces interaction between the views which allows information form all other cities to be taken into account during the training phase. The following is a brief discussion of MV LS-SVM and some results on temperature prediction. For a more in-depth discussion see [4].

2 Multi-View LS-SVM Regression (MV LS-SVM)

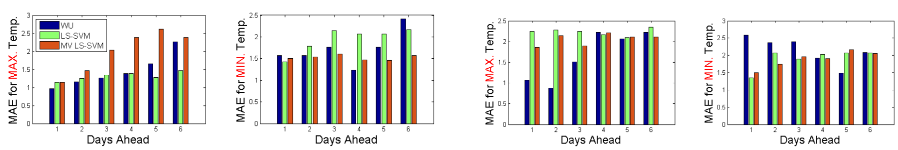
The primal formulation of MV LS-SVM consist of multiple LS-SVM regression objectives and a coupling term. This newly introduced term enforces the alignment of the error variables over multiple views so that when training on one view, the other views are taken into account. Given a number of V views, a training set of N data points $\{y_k, \mathbf{x}_k^{[v]}\}_{k=1}^N$ where $\mathbf{x}_k^{[v]} \in \mathbb{R}^{d^{[v]}}$ denotes the k-th input sample for view v and $y_k \in \mathbb{R}$ the k-th target value, the primal formulation of the proposed model is:

$$\begin{aligned} \min_{\substack{\mathbf{w}^{[v]}, \mathbf{e}^{[v]}, \\ b^{[v]}}} \ & \frac{1}{2} \sum_{v=1}^{V} \mathbf{w}^{[v]^T} \mathbf{w}^{[v]} + \frac{1}{2} \sum_{v=1}^{V} \gamma^{[v]} \mathbf{e}^{[v]^T} \mathbf{e}^{[v]} + \rho \sum_{v,u=1; v \neq u}^{V} \mathbf{e}^{[v]^T} \mathbf{e}^{[u]} \\ & \text{s.t. } \mathbf{y} = \boldsymbol{\Phi}^{[v]} \mathbf{w}^{[v]} + b^{[v]} \mathbf{1}_N + \mathbf{e}^{[v]} \quad \text{for } v = 1, \ldots, V \end{aligned} \tag{1}$$

where $\mathbf{y} = [y_1; \ldots; y_N]$ and $b^{[v]}$ are bias terms, $\gamma^{[v]}$ are positive real constants and $\mathbf{e}^{[v]} \in \mathbb{R}^N$ are error variables. $\boldsymbol{\Phi}^{[v]}$ is defined as $\boldsymbol{\Phi}^{[v]} = [\varphi^{[v]}(\mathbf{x}_1^{[v]})^T; \ldots; \varphi^{[v]}(\mathbf{x}_N^{[v]})^T]$ where $\varphi^{[v]} : \mathbb{R}^{d^{[v]}} \rightarrow \mathbb{R}^{d_h^{[v]}}$ are the mappings to a high dimensional feature space related to the vth view. By taking the Lagrangian of the primal problem, deriving the KKT optimality conditions and eliminating the primal variables, the dual problem, a linear system, is obtained.

3 Experiments and conclusion

The data have been collected from Weather Underground [5] and include real measurements for Brussels and 9 neighboring cities. It covers a time period from beginning 2007 to mid 2014 and comprises 198 measured weather variables for each day. The performance of the proposed method is evaluated on min. and max. temperature prediction in two test sets: (i) from mid-Nov. 2013 to mid-Dec. 2013 (Nov/Dec) and (ii) from mid-Apr. 2014 to mid-May 2014 (Apr/May).



(a) Nov/Dec

(b) Apr/May

Fig. 1: Comparison of the MAE on temperature prediction between Weather Underground (WU), LS-SVM on the concatenated features and MV LS-SVM.

Figure 1 summarizes the results. We can conclude that the proposed method MV LS-SVM can often outperform LS-SVM for min. and max. temperature prediction. It also shows that MV LS-SVM can outperform WU for min. temperature prediction and is competitive with it for max.temperature in the test set Apr/May. These results suggest the merit of using a multi-view approach instead of simply concatenating the features.

Acknowledgments

The research has received funding from the European Research Council under the European Union Seventh Framework Programme (FP7/2007-2013) / ERC AdG A-DATADRIVE- B (290923). This paper reflects only the authors views and the Union is not liable for any use that may be made of the contained information. Research Council KUL: CoE PFV/10/002 (OPTEC), BIL12/11T; PhD/Postdoc grants Flemish Government: FWO: projects: G.0377.12 (Structured systems), G.088114N (Tensor based data similarity), G0A4917N (Deep restricted kernel machines); PhD/Postdoc grant iMinds Medical Information Technologies SBO 2015 IWT: POM II SBO 100031 Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, Dynamical systems, control and optimization, 2012-2017).

References

1. Blum, A., Mitchell, T.: Combining labeled and unlabeled data with co-training. Conference on Learning Theory (1998) 92–100
2. Suykens, J.A.K., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J.: Least Squares Support Vector Machines. World Scientific (2002)
3. Lazo, J.K., Morss, R.E., Demuth, J.L.: 300 billion served: Sources, perceptions, uses, and values of weather forecasts. Bulletin of the American Meteorological Society **90**(6) (2009) 785–798
4. Houthuys, L., Karevan, Z., Suykens, J.A.K.: Multi-view LS-SVM regression for black-box temperature prediction in weather forecasting. Proceedings of the International Joint Conference on Neural Networks (IJCNN) (2017) 1102–1108
5. Weather Underground Co.: `www.wunderground.com`