Bleeding Detection by Multi-View Correlation Clustering of Central Venous Pressure

Eric Lei
Machine Learning Department
Carnegie Mellon University
elei@cs.cmu.edu

Kyle Miller
Auton Lab
Carnegie Mellon University
mille856@andrew.cmu.edu

Michael R. Pinsky
School of Medicine
University of Pittsburgh
pinsky@pitt.edu

Artur Dubrawski
Auton Lab
Carnegie Mellon University
awd@cs.cmu.edu

Abstract

One interesting question in medicine asks whether the presence of bleeding is correlated with central venous pressure (CVP), which is often used as a signal in practice. Although several authors argue that CVP is useless because it is noisy, the question of filtering noise is separate from the question of CVP's intrinsic utility. In this study we consider a dataset recorded in a controlled setting to reduce noise. The main contribution of this work is to demonstrate that these clean CVP waveforms can help predict bleeding. In particular, bleeding has a relationship with correlation structures between waveforms during inspiration and expiration phases of the respiratory cycle. We employ a multi-view correlation clustering method called CLS clustering that learns mixtures of local correlation structures to classify sets of waveforms as bleeding or not bleeding, with fair empirical performance. Clusters represent clinical phenotypes, and the relationship between inspiration and expiration can be expressed in terms of the original CVP waveforms.

1 Introduction

An area of interest in medicine is the detection of internal bleeding. One interesting question is whether the presence of bleeding is correlated with central venous pressure (CVP), the blood pressure in a region near the right atrium of the heart. Several authors argue that there is no empirical evidence that CVP has any clinical utility even though it is often used in practice (Marik and Cavallazzi, 2013). A commonly cited reason is that CVP is highly sensitive to shifts in body position, making it too noisy. However, the question of whether CVP is connected to bleeding is separate from the challenge of filtering out this noise. In this work we investigate CVP within a controlled setting that restricts this noise. To aid scientific discovery, we explore a method that reveals explainable structure in the data. We consider a recent classification method that learns a clustering of local linear relationships between two views in the data (Lei et al., 2017), and we apply it to the task of bleeding detection. In the dataset used by this study, the two views correspond to CVP waveforms during inspiration and expiration phases of the respiratory cycle.

The chosen method finds clusters of correlation structures between two multivariate views of the data. This approach is useful when different correlation structures appear in different subsets of the data and when nonlinear correlations may be present. Inspired by Canonical Correlation Analysis (Hotelling, 1936), the method is called Canonical Least Squares (CLS) clustering; CLS clusters can be considered shared factor models between the first and second views (Lei et al., 2017). The method supports must-

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

link constraints between data points, which are utilized in this study to guarantee that observations from the same patient appear in the same cluster. Clusters can be interpreted as clinical phenotypes characterizing patients' pre-bleeding or post-bleeding responses. This study utilizes CLS clustering as a supervised classification algorithm by incorporating labels on bleeding status. The result is a method with fair classification performance with interpretable structure.

2 Related work

There has been substantial past work on multi-view clustering. However, many authors, such as Livescu and Stoehr (2009) and Bruno and Marchand-Maillet (2009), work with multi-modal data such as audio-visual data; clinical data is less common. Furthermore, a large body of literature has been published on the clinical utility of CVP in practice. Most studies appear to post negative results, such as Michard and Teboul (2000), Pinsky and Payen (2005), and Kumar et al. (2004); some studies claim positive results, such as Damman et al. (2009) and Boyd et al. (2011). However, less work has been published on investigating CVP in a controlled, noiseless setting.

3 Canonical correlation analysis

CCA is a method for understanding cross-covariance between two sets of variables (Hotelling, 1936). For example, if one set is genes and the other is diseases, then CCA might connect combinations of genes with certain diseases, potentially corresponding to physiological traits. Formally, it finds maximally correlated linear combinations of each set. Let $x \in \mathbb{R}^{d_1}$ and $y \in \mathbb{R}^{d_2}$ be random vectors. CCA solves the problem $\max_{u,v} \operatorname{Corr}(x^\intercal u, y^\intercal v)$. The solution is well-understood (Hardoon et al., 2004) as the solution to an eigenvalue problem involving covariance matrices. Subsequent linear combinations can be found under the constraint that the previous components $x^\intercal u$ and $y^\intercal v$ are uncorrelated with the new ones.

4 Multi-view correlation clustering

We summarize a method for multi-view correlation clustering called Canonical Least Squares (CLS) clustering introduced by Lei et al. (2017). To analyze correlations between two views, we identify multi-view factor analysis, including CCA, as a useful tool. We consider a mixture of local multi-view correlation models. Formally, given a dataset described by two sets of variables x and y, a hyperparameter for the number k of clusters, and a hyperparameter for the number m of canonical variables to use, the goal of CLS clustering is to partition the data into k clusters such that for instances in the same cluster, the features in x and y are correlated in the same way. A natural choice for the correlation models would be CCA, which was explored by Fern and Friedl (2005); however, Lei et al. (2017) propose an alternative choice, CLS, with useful properties for clustering.

Like CCA, CLS takes views of data X and Y and produces m pairs of vectors (u, v) such that the components $X^\intercal u$ and $Y^\intercal v$ have some kind of relationship. Unlike CCA, this relationship is not of maximum correlation but of least squared error. Clustering via CLS has a few advantages over CCA. For example, the existence of a well-formulated objective is missing from CCA clustering. One practical benefit of this difference is that CLS clustering can be run with many different initializations, and the solutions can be compared by the objective function to select the best one. In addition, when $m = 1$, the objective function is guaranteed to never increase. Furthermore, this method permits must-link constraints that designate sets of points that must appear in the same cluster. These constraints can be encoded by assigning each set of points to the cluster that minimizes the sum of squared errors over the points.

The clustering algorithm employs a structure similar to the expectation-maximization algorithm (Dempster et al., 1977), alternating between update and assignment steps. Let $R^{(i)} \in \{0,1\}^{n\times n}$ be a diagonal matrix whose element at (j, j) indicates whether observation j is in cluster i.

Update step. To find m components given cluster assignments $R^{(i)}$ for each cluster i, the following optimization problem is solved: $\sum_i \min_{U^{(i)},V^{(i)}} \|R^{(i)}(XU^{(i)} - YV^{(i)})\|_{\mathcal{F}}^2$ subject to $V^{(i)\intercal}V^{(i)} = I,\ i = 1,\dots,k$, where $U^{(i)} \in \mathbb{R}^{d_1\times m}$ and $V^{(i)} \in \mathbb{R}^{d_2\times m}$ are the coefficients for X and Y respectively in cluster i.

Assignment step. To assign point (x, y) to a cluster given CLS coefficients U and V, find $\mathrm{argmin}_i \|(x^\top U^{(i)} - y^\top V^{(i)})\|_{\mathcal{F}}^2$.

5 Experiments

5.1 Dataset

The data were collected from an experiment in which pigs were subjected to controlled bleeding. The experimental procedure was similar to that in Pinsky (1984). Thirty-eight Yorkshire pigs were sedated and bled at a constant rate of 20 mL/min. Their central venous pressure (CVP), the blood pressure in a major vein to the heart, was monitored for 20 minutes before bleeding and 30 minutes after. Two CVP waveforms (Fig. 1) were extracted from each respiration cycle, one from the inspiration phase of breathing and the other from the expiration phase. Inspiration and expiration represent the two views of the data in our experiment. The respiration cycles lasted 5.2 seconds on average, resulting in an average of 556 cycles per pig over the 50 minutes. Twenty-one features were extracted from each waveform as averages and ratios between different points.

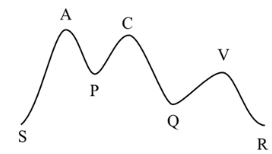


Figure 1: Central venous pressure waveform.

5.2 Procedure

This work employed CLS clustering as a supervised classification method by first partitioning the data into two sets according to bleeding status. For each of these sets, CLS clusters were learned. Each pig's observations were constrained to belong to the same cluster. Next, test observations were classified by finding the best fitting cluster in each of the pre-bleeding and post-bleeding models and then computing the difference in squared-error between those two best clusters as a score. The method was run 64 times with different random initializations, and the scores were averaged to form the final score. This score was then used to compute the area under the receiver operating characteristic curve (AUC), true positive rate (TPR) at a false positive rate (FPR) of 10%, and FPR at a TPR of 50%. This testing process was used in leave-one-out cross-validation on a training set of pigs to select the hyperparameters of numbers of clusters and components. We selected 5 clusters for pre-bleeding and 6 for post-bleeding as well as 4 components for both models.

On the test set of pigs, the performance metrics for each pig was computed by the above process, and the values were then averaged. In addition, we tried classifying consecutive windows of G observations from a given pig constraining them to the same cluster. In the above process, $G = 1$ because observations are classified individually, but when G is greater the score is less noisy because there are more points to make each classification. Also, we tried classifying the entire pre-bleeding or post-bleeding phase of each pig.

5.3 Results

Table 1: My caption

Window size	AUC	TPR@10FPR	FPR@50TPR
1	87.9 ± 5.3	72.8 ± 14.4	5.5 ± 5.5
12	89.4 ± 6.0	76.5 ± 16.9	4.7 ± 5.6
All observations	92.3 ± 15.4	-	-

The different configurations resulted in performance metrics given by Table 1, which in part summarize Fig. 2. As expected, when more observations were classified together, the classification performance improved. This trend is illustrated by Figs. 4a and 4b, in which the predictions for $G = 12$ are much smoother than for $G = 1$. When only one cluster was used for the pre-bleeding or post-bleeding models, the AUC was 55.4 ± 8.0, which showed the benefit of using a mixture of local models rather than a global model.

We examine latent factors that determine classification in Fig. 3. The factors, constructed as linear combinations of each view, are expected to align and have zero residual. This pattern holds before

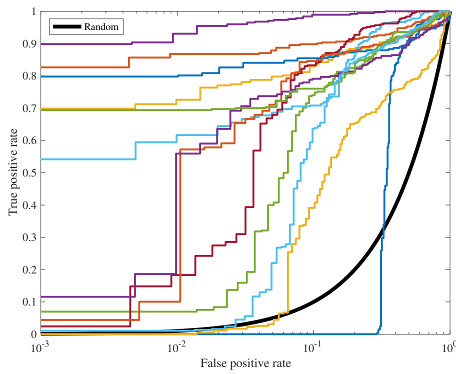


Figure 2: ROCs of pigs in the test set with $G = 1$.

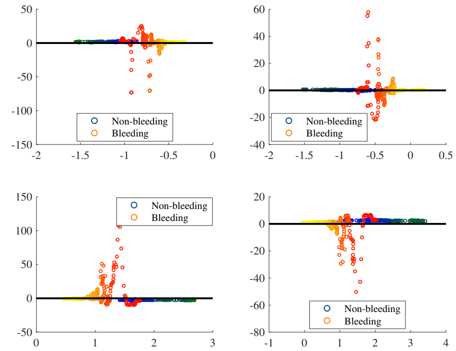
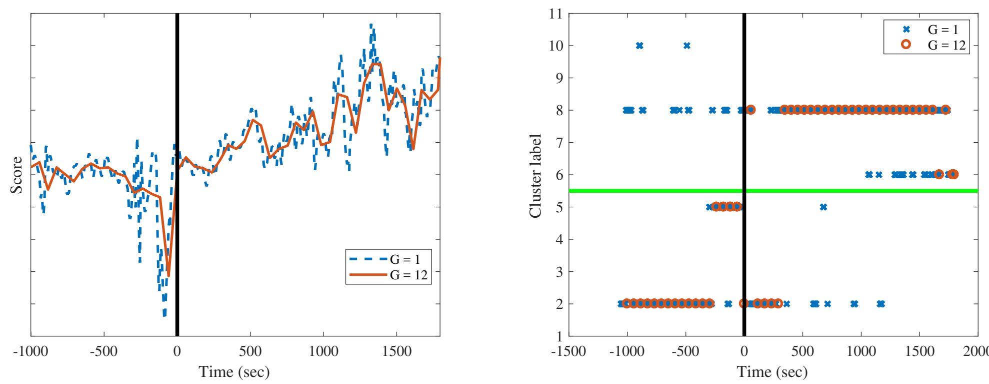


Figure 3: Latent variable residuals in Cluster 2 from one pig. Residuals diverge from 0 when bleeding starts.



(a) Bleeding classification scores.

(b) Cluster assignments. Clusters below the green line correspond to pre-bleeding.

Figure 4: Scores and cluster assignments over time from this approach for a single pig. Bleeding begins at $t = 0$.

bleeding but is violated after bleeding, which indicates that this cluster fits only before bleeding. By examining the coefficients in the linear combinations, it may be possible to find interesting clinical interpretations of the factors.

6 Conclusion

The main contribution of this work was to demonstrate that CVP waveforms can help predict bleeding when the level of noise is controlled. In particular, bleeding has a relationship with correlation structures between waveforms during inspiration and expiration. We employed a method called CLS clustering that learns mixtures of local correlation structures to train models for pre-bleeding and post-bleeding waveforms. The method clustered the pre-bleeding and post-bleeding phases of each pig. One type of interpretability it offered was that sets of new observations were not only classified by bleeding status but also assigned to clusters or phenotypes within their classes. In addition, the classification decisions were based on interpretable factors from each view.

Acknowledgements

This work has been partially funded by NIH award R01NR013912 and by DARPA award FA8750-17-2-0130.

References

Boyd, J. H., Forbes, J., Nakada, T.-a., Walley, K. R., and Russell, J. A. (2011). Fluid resuscitation in septic shock: a positive fluid balance and elevated central venous pressure are associated with increased mortality. *Critical care medicine*, 39(2):259–265.

Bruno, E. and Marchand-Maillet, S. (2009). Multiview clustering: A late fusion approach using latent models. In *ACM SIGIR Conference on Research and Development in Information Retrieval*, number January, page 736.

Damman, K., van Deursen, V. M., Navis, G., Voors, A. A., van Veldhuisen, D. J., and Hillege, H. L. (2009). Increased central venous pressure is associated with impaired renal function and mortality in a broad spectrum of patients with cardiovascular disease. *Journal of the American College of Cardiology*, 53(7):582–588.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.

Fern, X. and Friedl, M. (2005). Correlation clustering for learning mixtures of canonical correlation models. In *SIAM International Conference on Data Mining*, pages 439–446.

Hardoon, D. R., Szedmak, S., and Shawe-Taylor, J. (2004). Canonical correlation analysis: An overview with application to learning methods. *Neural Computation*, 16(12):2639–2664.

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28(3/4):321–377.

Kumar, A., Anel, R., Bunnell, E., Habet, K., Zanotti, S., Marshall, S., Neumann, A., Ali, A., Cheang, M., Kavinsky, C., et al. (2004). Pulmonary artery occlusion pressure and central venous pressure fail to predict ventricular filling volume, cardiac performance, or the response to volume infusion in normal subjects. *Critical care medicine*, 32(3):691–699.

Lei, E., Miller, K., and Dubrawski, A. (2017). Learning Mixtures of Multi-Output Regression Models by Correlation Clustering for Multi-View Data. In *eprint arXiv:stat.ML/1709.05602*.

Livescu, K. and Stoehr, M. (2009). Multi-view learning of acoustic features for speaker recognition. In *IEEE Workshop on Automatic Speech Recognition and Understanding*, pages 82–86.

Marik, P. E. and Cavallazzi, R. (2013). Does the central venous pressure predict fluid responsiveness? an updated meta-analysis and a plea for some common sense. *Critical care medicine*, 41(7):1774–1781.

Michard, F. and Teboul, J.-L. (2000). Using heart-lung interactions to assess fluid responsiveness during mechanical ventilation. *Critical Care*, 4(5):282.

Pinsky, M. R. (1984). Instantaneous venous return curves in an intact canine preparation. *Journal of Applied Physiology*, 56(3):765–771.

Pinsky, M. R. and Payen, D. (2005). Functional hemodynamic monitoring. *Critical Care*, 9(6):566.