

Improving Surgical Models through One/Two Class Learning

Chih-Chun Chia, Zahi Karam, Gyemin Lee, Ilan Rubinfeld, and Zeeshan Syed

Abstract—Only a minority of patients undergoing in-patient surgical procedures experience complications. However, the large number of in-patient surgeries (over 48 million procedures each year in the U.S.) results in substantial overall mortality and morbidity due to these complications. This burden can be decreased through improvements in the ability to evaluate patients by the bedside, and to assess surgical quality and outcomes across hospitals. Unfortunately, the process of developing clinical models for surgical complications is made challenging by the availability of generally small datasets for model training, and by class imbalance due to the diminished prevalence of many important complications. In this paper, we address this issue and explore the idea of jointly leveraging the benefits of both supervised and unsupervised learning to model surgical complications that occur infrequently. In particular, we study an approach where the problems of supervised and unsupervised model development are treated as tasks that can be transferred. Focussing this work on support vector machine (SVM) classification, we describe a transfer learning algorithm that improves performance relative to both supervised (i.e., binary or 2-class SVM) and unsupervised (i.e., 1-class SVM) methods, as well as the use of cost-sensitive weighting techniques, for predicting different surgical complications within the American College of Surgeons National Surgical Quality Improvement Program registry.

I. INTRODUCTION

Surgical models that can accurately discriminate between patients at high or low risk of complications following in-patient procedures are an important resource in reducing the mortality and morbidity associated with surgery. These models are valuable clinical tools in evaluating patients by the bedside, as well as to assess quality and outcomes across hospitals. However, predicting surgical complications has traditionally represented a challenging proposition, with no satisfactory predictors for many adverse clinical outcomes.

The challenges of developing models for surgical complications usually stem from existing datasets available for model derivation being small (e.g., thousands to tens of thousands of patients) and suffering from class imbalance. Traditionally, models to stratify surgical patients have been developed within a *supervised* learning framework. However, supervised learning approaches focus on characterizing the differences between patients who do or do not experience clinical events, and suffer from the lack of sufficient positive (i.e., event) examples for model training when clinical events occur with diminished prevalence. For example, the rate of a wide range of serious complications, ranging from coma

to bleeding requiring transfusion was well below 1% in the American College of Surgeons National Surgical Quality Improvement Program (NSQIP) data sampled at over 200 hospital sites [1]. Collecting additional data to address this issue of class imbalance during model training is often infeasible because of delays and expenses to both patients and caregivers. The costs and complexity of collecting extensive data annotated by experts have impeded the spread of well validated and effective healthcare quality interventions [2].

There is a growing body of recent work [3], [4] focusing on addressing these issues in the context of *unsupervised* learning. In the presence of small datasets with few positive examples, these efforts evaluate patients by learning the support of the available data, and by comparing the clinical characteristics of new patients to the distribution of existing patient records. In studies on different clinical applications, these approaches successfully discriminate patients at increased or decreased risk of clinical events. While these results are promising, however, in general the unsupervised learning approaches do not consistently improve performance relative to models developed through supervised learning.

In this study, we build upon these earlier results using unsupervised learning while developing models for surgical complications. We note that a limitation of earlier work on unsupervised learning in medicine is that it does not exploit any of the information available in labeled examples. While the absence of a large number of positive examples makes it difficult for supervised models to generalize, we nevertheless believe that there may be useful information in the patient labels beyond the support of these patients in the feature space that can be exploited. Most notably, a key limitation of unsupervised algorithms is that paradoxically they consider both the healthiest and unhealthiest individuals in a population as being at highest risk (since these examples are most likely to manifest as tails of the patient-population distribution). By using label information for patients, we believe that it may be possible to encode the directionality of these tails for surgical evaluation.

We exploit this observation and explore the idea of jointly leveraging the benefits of both supervised and unsupervised risk stratification of patients undergoing surgery. We do so by adopting an approach that can be considered as a specialized case of transfer learning. Transfer learning [5] is typically used to address situations where the data for model training and model application are drawn from different distributions. In such cases, transfer learning provides a way to refine a model between training and application. In our work, we adopt a different use of transfer learning: instead of using transfer learning to refine a model across datasets, we use the

*This work was supported by the National Science Foundation and the American Heart Association

C. Chia, Z. Karam, G. Lee and Z. Syed are with the University of Michigan, Ann Arbor, USA; I. Rubinfeld is with the Henry Ford Hospital, Detroit, USA.

same underlying principles to refine a model developed using supervised learning for use in unsupervised learning (and vice versa). In this way, our approach uses transfer learning to refine a model across tasks or problem formulations rather than datasets while modeling surgical complications.

Popular approaches for transfer learning include instance-based, feature-based, and parameter-based algorithms [5]. In instance-based transfer learning, the goal is to address the issue of training and test distribution differences by reweighting data in the source domain for use in target domain. Feature-based transfer learning finds a good feature representation that is common to both source and target domains. Since both instance-based and feature-based transfer learning are essentially focused on the situation of refining a model between datasets, they are not relevant to our work. Instead, the approach of parameter-based transfer learning, which focuses on transferring parameters between similar but distinct tasks is most relevant to the clinical problem considered here. We explore this idea in our work, and: (i) introduce the idea of combining the merits of both supervised and unsupervised learning for clinical risk stratification; (ii) propose a formulation that exploits this idea by describing a transfer learning-based algorithm that leverages information in both supervised and unsupervised tasks; (iii) evaluate our ideas and the use of our SVM-based algorithm in a rigorous investigation on a real-world population of surgical patients; and (iv) describe the strengths and limitations of our research as a means of informing future research.

II. BACKGROUND

We position our work on modeling surgical complications within the framework of support vector machine (SVM) classification, and here we present a background description of the 2-class and 1-class SVM.

A. 2-Class Support Vector Machine Classification

Binary or 2-class SVM [6] focuses on learning a hyperplane in a high-dimensional feature space that can be used for classification. Given a training set $\{(\mathbf{x}_i, y_i) | \mathbf{x}_i \in \mathbb{R}^m, y_i \in \{+1, -1\}\}_{i=1}^n$ the soft margin SVM formulation aims to solve the following constrained optimization problem:

$$\begin{aligned} \min_{\mathbf{w}, \xi} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & y_i(\mathbf{w}^T \phi(\mathbf{x}_i) - b) \geq 1 - \xi_i \quad \forall i = 1, \dots, n \\ & \xi_i \geq 0 \end{aligned}$$

where ϕ is a kernel function that maps data into some feature space, and the constant C reflects the cost of misclassification and the ξ_i correspond to the slack variables of the soft margin SVM. The dual form of the problem is given by:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbb{K}(i, j) - \sum_{i=1}^n \alpha_i \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq C \quad \forall i = 1, \dots, n \\ & \sum_{i=1}^n \alpha_i y_i = 0 \end{aligned}$$

where $\mathbb{K}(i, j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ is the kernel matrix [7]. The final classification rule for predicting the label of a new example \mathbf{x} is then given by $\hat{y} = \text{sgn}(\mathbf{w}^* \cdot \phi(\mathbf{x}) - b)$, and $\mathbf{w}^* = \sum_{i=1}^n \alpha_i y_i \phi(\mathbf{x}_i)$ can be obtained by solving the dual formulation.

B. 1-Class Support Vector Machine Classification

The 1-class SVM [8] aims to estimate the support S of a high-dimensional distribution such that the probability that a point drawn from the input space lies outside S is low. Roughly speaking, in contrast to the 2-class SVM algorithm, which separates two classes in the feature space by a hyperplane, the 1-class SVM attempts to separate the entire dataset from the origin. Given training data of the form $\{(\mathbf{x}_i) | \mathbf{x}_i \in \mathbb{R}^d\}_{i=1}^n$ (i.e., with the class labels either not available or ignored for training in an unsupervised setting), the 1-class SVM solves the following quadratic problem (which penalizes feature vectors not separated from the origin, while simultaneously trying to maximize the distance of this hyperplane from the origin):

$$\begin{aligned} \min_{\mathbf{w}, \psi, \rho} \quad & \frac{1}{2} \|\mathbf{w}\|^2 - \rho + C \sum_{i=1}^n \psi_i \\ \text{s.t.} \quad & \mathbf{w}^T \phi(\mathbf{x}_i) - \rho \geq (-\psi_i) \quad \forall i = 1, \dots, n \\ & \psi_i \geq 0 \end{aligned}$$

where the constant C expresses the tradeoff between incorporating outliers that are not separated from the origin and minimizing the support region. The dual form of the 1-class SVM problem is:

$$\begin{aligned} \min_{\alpha} \quad & \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \mathbb{K}(i, j) \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq C \quad \forall i = 1, \dots, n \\ & \sum_{i=1}^n \alpha_i = 1 \end{aligned}$$

The classification rule for predicting whether a new example \mathbf{x} lies within the region of high probability is then given by $\hat{y} = \text{sgn}(\mathbf{w}^* \cdot \phi(\mathbf{x}) - \rho)$ with $\hat{y} \leq 0$ denoting the detection of an outlier, and $\mathbf{w}^* = \sum_{i=1}^n \alpha_i \phi(\mathbf{x}_i)$ can be obtained by solving the dual formulation.

III. METHOD

We propose a transfer learning extension to 2-class and 1-class SVM classification. Given the training set $\{(\mathbf{x}_i, y_i) | \mathbf{x}_i \in \mathbb{R}^d, y_i \in \{+1, -1\}\}_{i=1}^n$, we first utilize 2-class SVM classification for finding a maximum margin boundary \mathbf{w}_2^* . Our transfer learning formulation then transfers this 2-class boundary to the 1-class SVM task by solving the following optimization problem:

$$\begin{aligned} \min_{\mathbf{w}, \psi, \rho} \quad & \frac{1}{2} \|\mathbf{w} - \mathbf{w}_2^*\|^2 - \rho + C \sum_{i=1}^n \psi_i \\ \text{s.t.} \quad & \mathbf{w}^T \phi(\mathbf{x}_i) - \rho \geq (-\psi_i) \quad \forall i = 1, \dots, n \\ & \psi_i \geq 0 \end{aligned}$$

This model regularizes the 1-class SVM solution \mathbf{w} towards the model parameter \mathbf{w}_2^* obtained from the 2-class SVM classification task instead of regularizing \mathbf{w} by itself. In this setting, the regularization term C expresses the tradeoff between slacks and the distance between the transferred model and original model. The model learned will generally be closer to the 2-class SVM task model parameter \mathbf{w}_2^* when C has small values.

Similar to other SVM formulations, solving the dual of this optimization problem is more convenient and provides the advantage of using the kernel trick. In the interest of space, we only present the dual formulation and omit the derivation process (which can be easily done by introducing the Lagrangian). Also, since we already know from section II-A that $\mathbf{w}_2^* = \sum_{i=1}^n \alpha_i^* y_i \phi(\mathbf{x}_i)$, thus the dual form of the 2-to-1 SVM transfer problem can be written as:

$$\begin{aligned} \min_{\alpha} \quad & \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \mathbb{K}(i, j) + \mathbf{w}_2^* \sum_{i=1}^n \alpha_i \phi(\mathbf{x}_i) \\ & = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \mathbb{K}(i, j) + \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j^* y_j \mathbb{K}(i, j) \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq C \quad \forall i = 1, \dots, n \\ & \sum_{i=1}^n \alpha_i = 1 \end{aligned}$$

The transfer learning algorithm can also be applied to first find the one-class SVM boundary \mathbf{w}_1^* , and then transfer it using 2-class SVM. The primal formulation is given here:

$$\begin{aligned} \min_{\mathbf{w}, \xi} \quad & \frac{1}{2} \|\mathbf{w} - \mathbf{w}_1^*\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & y_i (\mathbf{w}^T \mathbf{x}_i - b) \geq 1 - \xi_i \quad \forall i = 1, \dots, n \\ & \xi_i \geq 0 \end{aligned}$$

We omit the interpretation here since it follows naturally from the previous discussion when transferring from two-class SVM to one-class SVM. Since we already know from section II-B that $\mathbf{w}_1^* = \sum_{i=1}^n \alpha_i^* \phi(\mathbf{x}_i)$, the dual form of the 1-to-2 SVM transfer problem is given by:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbb{K}(i, j) + \mathbf{w}_1^* \sum_{i=1}^n \alpha_i y_i \phi(\mathbf{x}_i) - \sum_{i=1}^n \alpha_i \\ & = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbb{K}(i, j) + \sum_{i=1}^n \alpha_i \left(\sum_{j=1}^n \alpha_j^* y_j \mathbb{K}(i, j) - 1 \right) \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq C \quad \forall i = 1, \dots, n \\ & \sum_{i=1}^n \alpha_i y_i = 0 \end{aligned}$$

IV. EVALUATION METHODOLOGY

A. Data

Our dataset comprised records from the American College of Surgery National Surgical Quality Improvements Program

(ACS-NSQIP) for patients undergoing surgery from 2005-2008 [1]. This data was used to develop models for each of the ten morbidity outcomes with lowest prevalence within 30 days of in-patient surgery (coma > 24 hours, peripheral nerve injury, myocardial infarction, stroke or cerebrovascular accident, pulmonary embolism, failure of extracardiac graft or prosthesis, renal insufficiency, cardiac arrest, renal failure, and bleeding requiring transfusion). After removal of cases with missing covariates, 27,673 patients data from year 2007 were used for training and 21,204 patients data from year 2005-2006 were used for internal selection of model parameters during derivation. NSQIP data for 2008 was used to test models on 35,657 patients.

B. Evaluation

For a given SVM model, each patient was assigned a score defined as the distance of the patient's feature vector from the decision boundary. We assessed the predictive ability of the SVM models by calculating the area under the receiver operating characteristic curve (AUROC) for the test patient scores relative to the different endpoints. Our choice of the AUROC as an evaluation metric is based on its widespread use to assess risk stratification models [9].

All algorithms were implemented using the MOSEK optimization toolbox for Matlab. Each model was trained using a linear kernel with cost parameters chosen by testing on the validation data from the set $\{C = 2^a | a \in [-10, 10]\}$. In addition to comparing the transferred SVM approaches to 2-class and 1-class SVM models, we also compared it to another popular approach that augments 2-class SVM model to address the class-imbalance issue [6]. In this method, the cost parameters were chosen to assign a weight to positive examples that was inversely proportional to how rarely they occurred in the data.

V. RESULTS

Table I presents two sets of comparisons. First, the last two columns of Table I compare the AUROC values obtained for the different surgical complications when transferring from 1-class to 2-class, and when transferring from 2-class to 1-class. For some endpoints, transferring from 1-class to 2-class outperformed the other approach, and vice versa for other endpoints. We note that this variation was not dependent on any obvious statistical property of the data (e.g., prevalence of complications). However, the improvement between the 1-to-2 or 2-to-1 transfer cases was consistent and could be identified within the training cohort in all cases. We therefore also consider an integrated model that is based on using the best of the 1-to-2 and 2-to-1 models (determined during training). Data for this approach is presented in Table I in the 'Combined transfer' column (third from the left).

The second comparison in Table I is between the transfer learning approaches and logistic regression (commonly used in clinical decision making and denoted by LR) [10], 1-class SVM, 2-class SVM, and 2-class SVM classification with

TABLE I

COMPARISON OF LOGISTIC REGRESSION, 1-CLASS SVM, 2-CLASS SVM, 2-CLASS SVM WITH RE-WEIGHTING, TRANSFERING FROM 1-CLASS TO 2-CLASS, TRANSFERING FROM 2-CLASS TO 1-CLASS, AND ALSO THE COMBINED TRANSFER MODEL AUROC VALUES FOR DIFFERENT ADVERSE OUTCOMES FROM THE NSQIP DATASET.

Endpoint	Occurrence in Test Set	LR	1-class SVM	2-Class SVM	2-Class SVM-W	Combined transfer	1 to 2 transfer	2 to 1 transfer
Coma	28 (0.1%)	0.615	0.700	0.601	0.560	0.759	0.601	0.759
Nerve Injury	56 (0.2%)	0.546	0.643	0.54	0.558	0.663	0.539	0.663
Myocardial Infarction	96 (0.3%)	0.746	0.694	0.741	0.729	0.774	0.774	0.709
Graft Failure	149 (0.4%)	0.772	0.704	0.770	0.748	0.781	0.781	0.705
Cerebrovascular Accident	173 (0.5%)	0.734	0.744	0.713	0.751	0.739	0.739	0.744
Pulmonary Embolism	223 (0.6%)	0.661	0.646	0.649	0.642	0.663	0.663	0.642
Renal Insufficiency	242 (0.7%)	0.731	0.732	0.703	0.733	0.744	0.703	0.744
Bleeding Event	284 (0.8%)	0.794	0.744	0.787	0.788	0.764	0.787	0.764
Cardiac Arrest	286 (0.8%)	0.805	0.777	0.783	0.802	0.805	0.783	0.805
Renal Failure	286 (0.8%)	0.831	0.773	0.807	0.821	0.812	0.812	0.791
Total Average		0.724	0.716	0.709	0.713	0.750	0.712	0.733

cost-sensitive weighting. Despite variations in the frequencies of outcomes and the numbers of patients in the training set, the combined transfer approach outperformed all the other methods in 7 out of the total 10 endpoints of interests. The total average AUROC was also higher for the transfer approach (0.750) than for LR, 1-class SVM, 2-class SVM, and 2-class SVM with cost-sensitive weighting (a maximum AUROC of 0.724 for any of the other approaches).

VI. DISCUSSION

In this study, we address the need for clinical models to stratify patients for surgical complications. We build upon previous work suggesting that in the absence of enough labeled training examples, an unsupervised learning approach may have value in stratifying patients for surgical events. Our research addresses an important potential limitation of this work, and describes how a purely unsupervised learning approach for developing clinical models can be improved by exploiting supplementary information available through supervised learning.

In this way, our approach resembles a transfer learning formulation for the goal of modeling surgical complications, with the unsupervised and supervised models framed as similar but distinct tasks. We are careful to distinguish our approach from transfer learning applications where the source and target domain data belong to different distributions. Instead, we propose using labeled and unlabeled versions of the same data for model transfer. To achieve this, we describe SVM-based algorithms that transfer the parameter of one model to another.

When evaluated on data from over 30,000 patients undergoing inpatient surgical procedures, our transfer SVM algorithm generally achieved better discrimination of patients at high risk of different morbidity outcomes than both 2-class and 1-class SVM models. In addition, our approach consistently outperformed 2-class SVM models where cost-sensitive weighting is used to overcome class imbalance, as well as the use of logistic regression.

We conclude with a brief discussion of some opportunities for future research. First, we believe that the question of how

the relative merits of transferring from 1-class to 2-class and vice versa change as the size of the training population and the prevalence of adverse outcomes varies is an interesting question that warrants further investigation. This information may be valuable in guiding the choice of the approach best suited for risk modeling in different clinical contexts (rather than selecting the best model in a purely empirical manner as was done in this study). Second, we propose exploring an approach that integrates the cost-sensitive weighting method for imbalanced learning with our transfer learning approach. Finally, we believe that there is a need to evaluate the ideas presented here in other clinical domains, and also relative to a broader set of ideas for addressing the issue of class imbalance. This may allow for a more complete appreciation of the contributions of our study.

REFERENCES

- [1] S.F. Khuri, The NSQIP: a new frontier in surgery, *Surgery*, 138(5):837-843, 2005
- [2] Schilling PL, Dimick JB, Birkmeyer JD. Prioritizing quality improvement in general surgery. *Journal of the American College of Surgeons*, 207(5):698-704, 2008
- [3] Syed Z, Rubinfeld I, Unsupervised risk stratification in clinical datasets: identifying patients at risk of rare outcomes, *International Conference in Machine Learning*. 1-8, 2010
- [4] Syed Z, Guttat J, Identifying patients at risk of major adverse cardiovascular events using symbolic mismatch. *Neural Information Processing Systems*. 1-9, 2010
- [5] Sinno Jialin Pan, Qiang Yang, A Survey on Transfer Learning, *IEEE Transactions on Knowledge and Data Engineering*, vol.22, no.10, pp.1345-1359, Oct. 2010
- [6] V. Vapnik. *Statistical Learning Theory*. Wiley, New York, NY, 1998. ch. 10-9
- [7] C.J.C. Burges, *Advances in kernel methods: support vector learning*, The MIT press, 1999
- [8] Scholkopf B, Platt J, Shawe-Taylor J, Smola A, Williamson R, Estimating the support of a high dimensional distribution, *Neural Computation*, 13:1443-1471, 2001
- [9] Altman DG, *Practical Statistics for Medical Research*, Chapman & Hall, 1991
- [10] Verplancke T, S Van Looy, Benoit D, Vansteelandt S, Depuydt P, De Turck F, Decruyenaer J, Support vector machine versus logistic regression modeling for prediction of hospital mortality in critically ill patients with haematological malignancies, *BMC Med Inform Dec Making* 8:56, 2008