Introduction to Machine Learning

Capstone Results and Final Report

Overview

Sepsis is the leading cause of mortality in the United States and the most expensive condition associated with in-hospital stays, accounting for 6.2% (nearly $24 billion) of total hospital costs. In particular, *Septic shock*, the most advanced complication of sepsis due to severe abnormalities of circulation and/or cellular metabolism, reaches a mortality rate as high as 50% and the annual incidence keeps rising. It is estimated that as many as 80% of sepsis deaths could be prevented with early diagnosis and intervention. Indeed, prior studies have demonstrated that *early diagnosis* and treatment of septic shock can significantly decrease patients' mortality and shorten their length of stay. In this capstone, the task is to build a machine learning model for accurate early diagnosis of septic shock.

Machine Learning Methods Compared

- Recent Temporal Pattern (RTP)
    - Support Vector Machine (SVM)
    - Logistic Regression (LR)
- Long Short-Term Memory (LSTM)

Methods with the highest metric scores

	Method	Metric	Score
0	RTP-svm	accuracy	0.842
1	RTP-lr	precision	0.831
2	RTP-svm	recall	0.977
3	RTP-svm	f1 score	0.860
4	RTP-lr	auc roc	0.917

Metrics Evaluation

- True Positive Rate (TP) - *The probability that an actual positive correctly tests positive.*
- True Negative Rate (TN) - *The probability that an actual negative correctly tests negative.*
- False Positive Rate (FP) - *The probability that an actual negative correctly tests negative.*
- False Negative Rate (FN) - *The probability that an actual positive wrongly tests negative.*

Accuracy: (TP + TN) / (TP + FP + FN + TN)

RTP-SVM correctly predicted a combined 84% of positive and negative cases. Not very informative for sepsis diagnosis due to the fact that false negatives (sepsis patients misidentified as non-sepsis patients) have a much higher cost than false positives (non-sepsis patients misidentified as sepsis patients) and Accuracy doesn't distinguish between the two.

Precision: TP / (TP + FP)

RTP-LR correctly classified 83% of its positive predictions. This tells us that our sepsis-positive predictions are reasonably reliable as long as the counter-measures for sepsis patients are not particularly invasive or resource intensive since misidentified non-sepsis patients would receive the same treatment as correctly identified sepsis patients.

Recall: TP / (TP + FN)

RTP-SVM correctly classified 97% of all positive cases. This is the most important metric for sepsis diagnosis since the accurate diagnosis of sepsis-positive patients is the priority here. As such with a 97% Recall rate we are reasonably certain of our ability to intercede on behalf of sepsis patients when necessary.

F1 Score: 2 x (Recall x Precision) / (Recall + Precision)

RTP-SVM has a fairly high 0.86 weighted average score between Recall and Precision. This is the ability to both identify sepsis cases (Recall) and be accurate with the cases it predicts as sepsis (Precision). In other words, the confidence we have in our ability to intervene on behalf of and spend resources on those patients, and only those patients, who truly need intervention.

AUC-ROC: Sensitivity (TP) vs (1 — Specificity (TN))

RTP-LR has the highest AUC-ROC score of 0.91 indicating its superior performance over LSTM at distinguishing between patients with sepsis and patients without sepsis at various thresholds.

Summary

Having completed a rigorous exploration of the sepsis EHR data using the methods RTP-SVM, RTP-LR, and LSTM, we have found that RTP outperformed LSTM on all of the relevant metrics. Further, the fitted models, SVM and LR, used in conjunction with RTP, each showed promise by achieving high scores in the various metrics. Of these, it is our conclusion that RTP-SVM is the most promising method due to the fact that it claims the high scores for the most relevant metrics for sepsis prediction, the Recall and F1-score. Using these scores, we can be moderately confident in our ability to intervene on behalf of sepsis patients while keeping unnecessary intervention levels to a reasonable minimum.

Proposal

After reviewing Dr. Chi's paper, "Recent Temporal Pattern Mining for Septic Shock Early Prediction", we found that their work cited an almost identical RTP F1-score of 0.868 using an observation window just 4 hours before diagnosis versus our F1-score of 0.860 using the same method 24 hours before diagnosis. That is a full 20 hours increase in advance notice with little drop-off in predictive ability. With such a minimal decrease in accuracy over such an extended period of time, we feel it would be reasonable to set a minimum acceptable F1 score and walk back the observation window an additional 24-48 hours in 4-hour intervals until the minimum F1 score is met. This will help to identify patients with sepsis that much earlier and before the disease has a chance to progress.