Deep Learning Based Sepsis Intervention: The Modelling and Prediction of Severe Sepsis Onset

Gavin Tsang,

XiangHua Xie

SwanseaVision
Swansea University • United Kingdom
csvision.swansea.ac.uk

Introduction

Sepsis an Overview – Epidemiology, Risk Factors, Prognosis

Introduction - Sepsis

- Life-threatening condition arising from the body’s response to infection.
- Diagnosis & Treatment:
  - Diagnosis:
    - qSOFA score – sepsis positive at 2/3 positively identified categories
      - Low blood pressure (SBP ≤ 100mmHg)
      - High respiratory rate (≥ 22 breaths/min)
      - Level of consciousness (GCS ≤ 14)
  - Treatment:
    - “Sepsis Six” – Antibiotics (<1H), Blood cultures, lactate & haemoglobin determination, urine output monitoring, high-flow oxygen, and intravenous fluids

Introduction - Sepsis

- Epidemiology:
  - 27.1% of adult ICU admission met severe sepsis criteria within 24H in the UK
  - Sepsis contributes 45% of ICU bed days & 33% of hospital bed days
  - Costs $24 billion annually in US (13% of US healthcare annual expenses)
- Prognosis:
  - 35% mortality rate before ICU discharge, 47% mortality rate during hospital spell
  - 63% hospital readmission within 1st year
  - **~4-8% increase in mortality rate per hour of delayed treatment**

Time critical prediction of sepsis

Related Work

Machine Learning based approaches to Sepsis prediction

Related Work

- Literary Review
  - Relevant papers from 2016 - 2019
- Focus on adult ICU patients
  - Ubiquity of high-frequency, detailed patient records
- Data characteristics
  - Features falling into patient vital signs, laboratory indications and demographic information
  - Populations ranging from 140 - 32,000 patients
- Machine learning applications
  - Varied modelling methodologies
    - Logistic Regression, SVM, Neural Network, Decision Trees
    - Most popular, Insight platform: ML based online decision support platform

Dataset & Objective

PhysioNet CinC 2019 Challenge

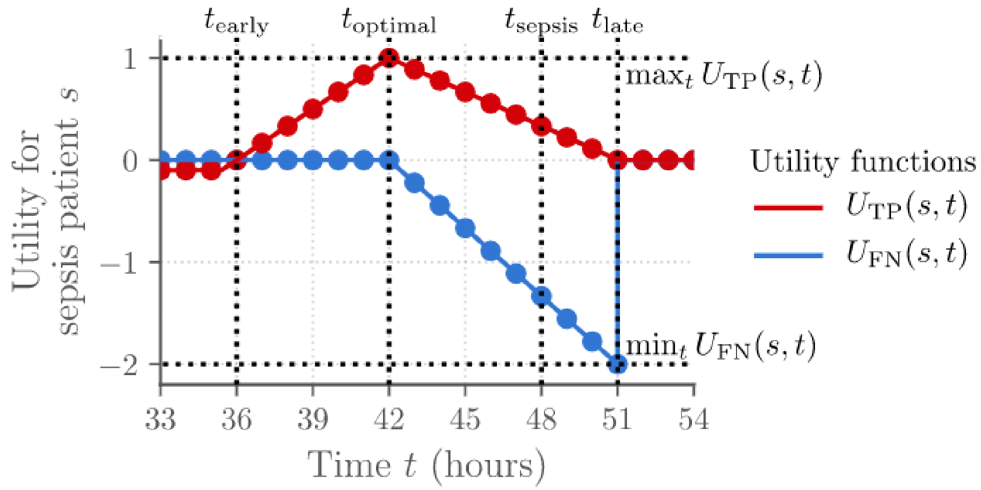
SwanseaVision
Swansea University • United Kingdom
csvision.swansea.ac.uk

Dataset – PhysioNet CinC 2019 Challenge

- ICU patient records from two hospital systems
  - A: Beth Israel Deaconess Medical Centre
  - B: Emory University Hospital
- 40 unique features
  - Categorised into vital signs, laboratory values, demographics
  - Highly sparse:
    - Vital signs: 32.4% missing values, Lab values: 94.9% missing values
- Hourly snapshots of 40,336 patients (55.9% male, 44.1% female)
  - Average age: 61.6 years, Standard Deviation: 16.5 years
  - 7.3% sepsis prevalence rate across dataset

Objective

- Hourly indications of positive/negative sepsis
  - Based on recorded clinical suspicion by medical practitioners
- Early prediction of sepsis development
  - 6 hours prior to recorded clinical suspicion
  - Hourly classification objective (+ve sepsis, -ve non-sepsis)
- Direct competition against previous challenge participants
  - Custom evaluation score – Utility score
  - Closeness to optimal -6H prediction point



Methodology

Boosted Cascading LSTMs

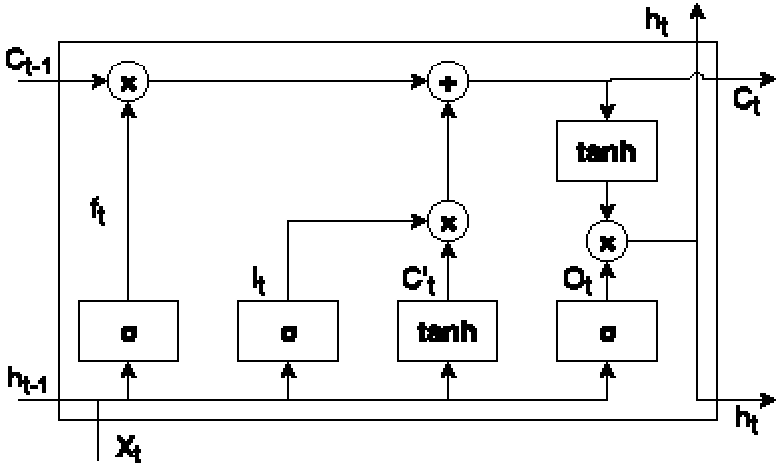
Shifting Margin Hinge Loss

Methodology - Overview

- Objective challenges
  - Complex non-linear associations between patient events and sepsis development
  - Mixture of continuous features and highly sparse features
  - High class imbalance between sepsis positive/negative timesteps
- Novel methodology
  - Augmenting tried and tested LSTM deep learning model
    - Boosted Cascading Sub-Networks
    - Shifting Margin Hinge Loss
    - Critical Diagnosis Point Penalty
    - Negative Reversal Penalty

Methodology – LSTM

- Quick Overview
  - Recurrent, time-shifted connection within LSTM node between sample timesteps
  - Excels in time-series based applications
    - Memory between timesteps
  - Formed of 4 components
    - Update gate
    - Forget gate
    - Input gate
    - Output gate



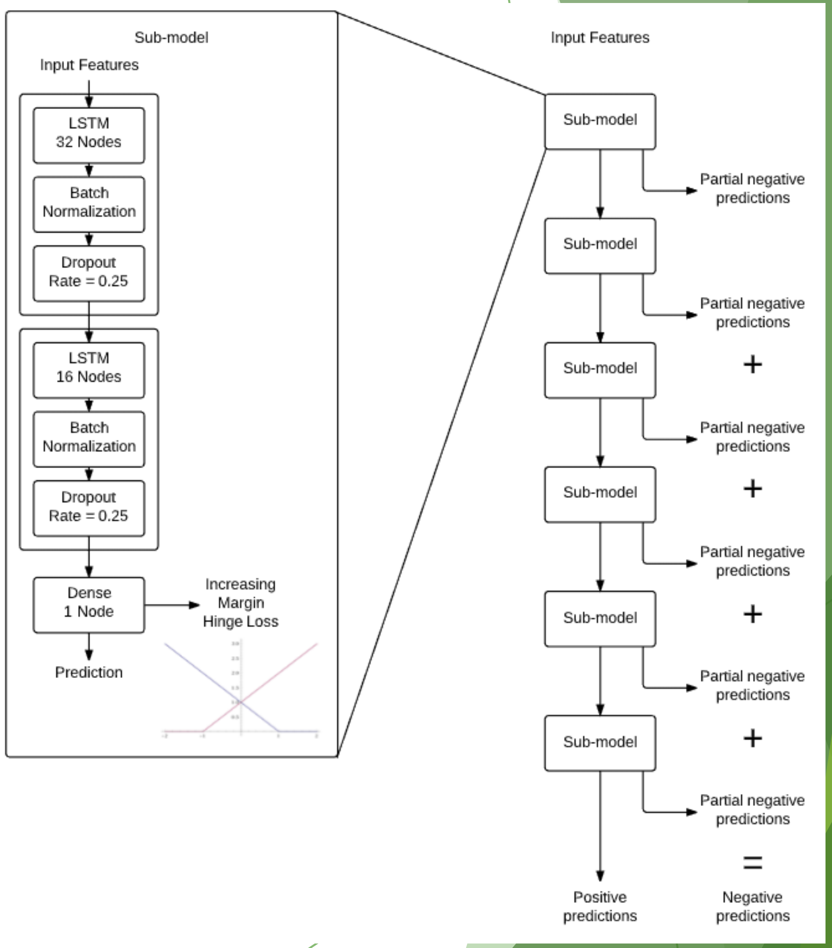
SwanseaVision
Swansea University • United Kingdom
csvision.swansea.ac.uk

Methodology – Boosted Cascading Sub-networks

- Cascading sub-models
  - Removal of confident negative predictions from large negative class subset
  - Lower cascades classify increasingly harder edge-case samples
  - Increasing model parameter capacity per cascade to allow for more complex distinction
- Boosted sampling
  - Emphasis on misclassified samples
  - Adaptive linear weighting of samples based on previous model classification

$$w_i^m = (1 - \lambda_w)w_i^{m-1} + \frac{\lambda_w}{T}\sum_{t=1}^{T}\left|y_{t,i} - \hat{y}_{t,i}^{m-1}\right|$$

- Improved performance on highly imbalanced datasets
  - Risk of over-fitting from smaller, less diverse dataset & larger model capacity
  - Shifting Margin Hinge Loss

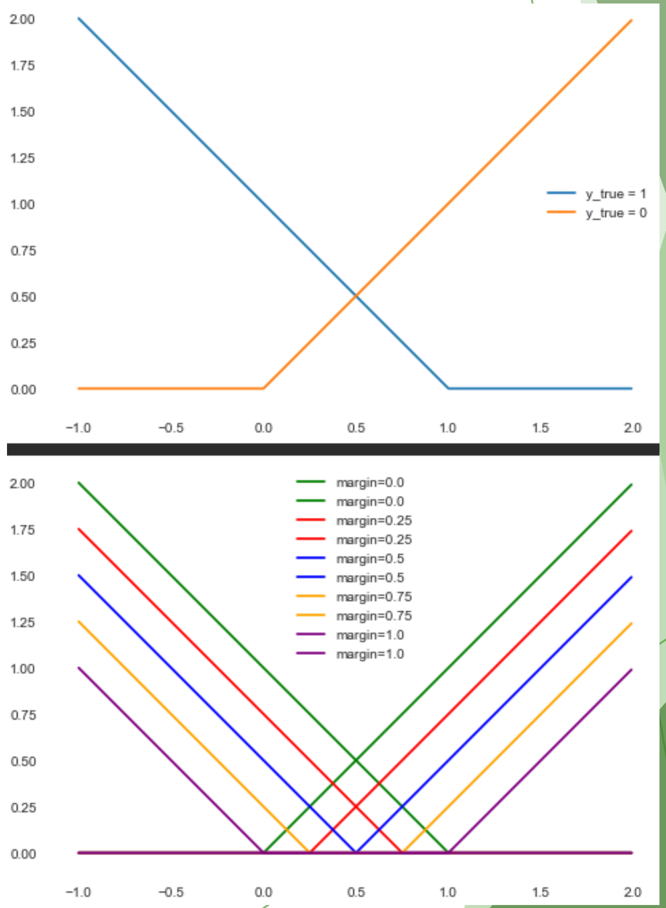


Methodology – Shifting Margin Hinge Loss

- Increasing model capacity & decreasing dataset size/diversity
  - Significant risk of over-fitting
- Shifting Margin Hinge Loss

$$L = \sum_{i=1}^{N} \max(0, \lambda_m - y_i \hat{y}_i)$$

- Adaptive selection of margin size within linear hinge loss
  - Margin controlled by defined super-parameter, $\lambda_m \in \{\mathbb{R} \geq 0\}$
  - Drives size of separating margin within feature space
  - Adjusts complexity of model's class separation hyperplane through larger/smaller margin



Methodology – Critical Diagnosis Point & Negative Reversal Penalty

- Hourly sepsis classification objective
  - Translates into regression objective of ICU stay duration till development of sepsis
  - Emphasis on critical $t_{optimal}$ diagnosis point indicating sepsis development
- Critical Diagnosis Point Penalty
  - Linear penalty regularisation function
  - Penalise early/late initial positive prediction of sepsis within timestep timeline.
  - Gradual penalisation allows for shifting of critical diagnosis point within training
- Negative Reversal Penalty
  - Sepsis development requires warning & physical intervention by medical practitioner
  - Heavily penalise reversion from previously positive sepsis prediction to a negative prediction in patient timeline

$$C_N = \lambda_N \sum_{t=0}^{N(\hat{y}_t)} \begin{cases} t' - t, & if\ \hat{y}_t = 0 \ni \hat{y}_{t'} = 1 \in \{\forall \hat{y}_{t'} : t' < t\} \\ 0, & otherwise \end{cases}$$

$$C_C = \lambda_C \sum_{t=0}^{N(\hat{y}_t)} \begin{cases} C_{TP}(t - t_{sepsis}), & \text{if } \hat{y}_t \text{ is TP} \\ C_{FN}(t - t_{sepsis}), & \text{if } \hat{y}_t \text{ is FN} \end{cases}$$

$$C_{TP}(t) = \begin{cases} \lambda_{TP} + \lambda_e, & \text{if } t < m_1(\lambda_{TP} + \lambda_e) + b_1 \\ m_1(t) + b_1, & \text{else if } t < t_{opt} \\ m_2(t) + b_2, & \text{else if } t < t_{late} \\ \lambda_{TP}, & \text{otherwise} \end{cases}$$

$$C_{FN}(t) = \begin{cases} \lambda_{TP}, & \text{if } t < t_{opt} \\ m_3(t) + b_3, & \text{else if } t < t_{late} \\ 1, & \text{otherwise} \end{cases}$$

where

$$m_1 = \frac{-\lambda_{TP}}{(t_{opt} - t_{early})}, \qquad b_1 = -m_1 t_{opt},$$

$$m_2 = \frac{\lambda_{TP}}{(t_{late} - t_{opt})}, \qquad b_2 = -m_2 t_{opt},$$

$$m_3 = \frac{1 - \lambda_{TP}}{(t_{late} - t_{opt})}, \qquad b_3 = -m_3 t_{late} + 1$$

Results

Comparison against PhysioNet CinC 2019 challenge participants

Results – Overall Prediction Metrics

Metric	Set A	Set B	Set A&B
True Pos. Rate	0.480±0.093	0.533±0.006	0.470±0.105
True Neg. Rate	0.982±0.010	0.985±0.002	0.977±0.019
False Pos. Rate	0.018±0.010	0.015±0.002	0.023±0.019
False Neg. Rate	0.520±0.093	0.467±0.006	0.530±0.105
Pos. Predictive Value	0.374±0.092	0.336±0.038	0.341±0.130
Neg. Predictive Value	0.988±0.002	0.993±0.000	0.990±0.003
False Omission Rate	0.012±0.002	0.007±0.000	0.010±0.003
False Discovery Rate	0.626±0.092	0.664±0.038	0.659±0.130
Accuracy	0.971±0.008	0.979±0.002	0.968±0.017
F1 Score	0.420±0.008	0.412±0.021	0.363±0.058
AUROC	0.855±0.032	0.893±0.026	0.737±0.142
AUPRC	0.391±0.010	0.351±0.042	0.258±0.051

- Experimental procedure:
  - Set A & B corresponding to separate hospital datasets
  - Set A test results use Set B as training set and *vice versa*. 5 repetitions performed for error margins
  - Set A&B test results use patient randomised 5 k-fold cross validation
- Results take-aways
  - Dataset is non-trivial
  - Significant improvement still available in model sensitivity
  - Issues in large class imbalance still apparent
    - Reflected in combined k-fold validation results
    - Significant increase in standard deviation due to variation in class balance within each fold

Results – Comparison against PhysioNet CinC 2019 challenge participants

Team Name	Utility Score		AUROC		AUPRC		Accuracy		F1 Score	
	Set A	Set B	Set A	Set B	Set A	Set B	Set A	Set B	Set A	Set B
Proposed Methodology	0.415	**0.450**	**0.855**	**0.893**	**0.391**	**0.351**	**0.971**	**0.979**	**0.420**	**0.412**
Can I get your signature?	**0.433**	0.434	0.000	0.000	0.000	0.000	0.828	0.888	0.139	0.140
Sepsyd	0.409	0.396	0.811	0.853	0.105	0.119	0.819	0.901	0.131	0.142
Separatrix	0.422	0.395	0.814	0.844	0.102	0.110	0.803	0.882	0.128	0.130
FlyingBubble	0.420	0.401	0.813	0.855	0.108	0.117	0.798	0.878	0.126	0.129
CTL-Team	0.401	0.407	0.806	0.846	0.101	0.116	0.797	0.891	0.122	0.137

- Results take-aways
  - All teams struggled with dataset class imbalance
  - Significant improvement in all categories except Set A utility score
  - Improvements to model sensitivity without sacrifice to model specificity provides such significant improvements to overall results
  - Utility score performance does not necessarily correlate to standard evaluation metric performance
  - Top performing team did not provide AUROC or AUPRC scores

Conclusion

Concluding remarks

Conclusion

- Objective:
  - Prediction of sepsis development at least 6 hours prior to official clinical suspicion
- Highly novel deep learning based methodology introduced:
  - Boosted Cascading Sub-Networks (LSTMs)
    - Provides effective prediction of highly imbalanced class ratios within time-series based data
  - Shifting Margin Hinge Loss
    - Provides effective adaptive regularisation of resulting over-fitting issues from said cascade.
  - Critical Diagnosis Point & Negative Reversal Penalty
    - Provides specialised penalty-based regularisation to emphasise sepsis prediction objective within a time-series based classification task

Conclusion

- Methodology drawbacks & potential future avenues
  - Significant improvements still available in model sensitivity
  - Use of two similar demographic datasets lack population scope
    - Evaluation on several alternative ICU based datasets (SAIL, MIMIC-III, etc.)
  - Adaption of methodology towards alternative modelling objectives
    - Model robustness in class imbalance allows for adaption towards other similar non-trivial patient medical record based objectives
- Overall
  - Proposed methodology significantly improves upon current state-of-the-art modelling techniques currently applied within medical informatics for prediction of sepsis onset within patients in an ICU setting via continuous patient medical records.