Comparing NLP Methods to Understand Clinical Text to Improve Outcomes in Septic patients

Gowri Nayar

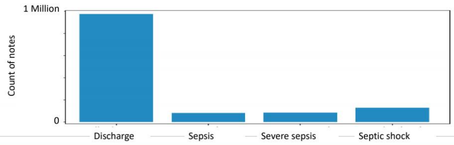
PhD Candidate in Biomedical informatics

Problem statement

- Patients at risk for sepsis are given a SOFA score to determine risk level and degree
- SOFA score is slow, dependent on vitals
- Can utilize nurse and clinician notes through NLP techniques

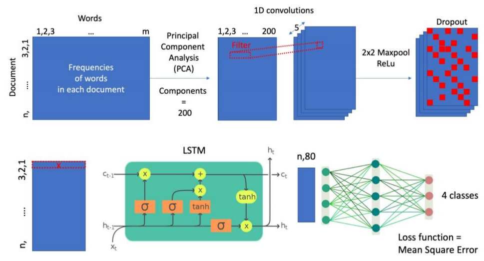
Background

- Clinical notes are seen as unstructured data that is difficult to use for automation within the biomedical community
- The increase of NLP tools that are pre-trained for medical applications can dramatically increase the utility of clinical notes
- Data is from critically ill visits, excluding patients who were diagnosed within the first 3 hours
- More than 40,000 patients were included with a large imbalance towards notes without a sepsis diagnosis, needs to be accounted for during training

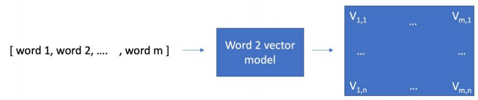


Methods

- Pre-process clinical notes to remove stop words, punctuation, and numbers
- In both methods, remove the high frequency and low frequency words
- Trained two networks for comparison, one based on a continuous bag of words method and the second on a Word2Vec embedding

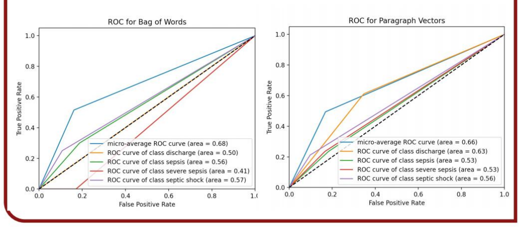


- Paragraph Vector Embeddding



Experiments

- Using the clinical notes and the models described above, we train and test each model using a 70/30 split for training and testing data (learning rate 0.01)
- To combat the bias in data, we use upsampling, by probabilistically augmenting the data with perturbed sepsis samples.
- Each model is compared against the true patient outcome and the clinically predicted outcome (SOFA score method)
- Both models outperform SOFA score, with accuracy of



Analysis

- Bag of Words model performs well on recall, particularly in the cases where there are distinctive, infrequent words, as in the case with Sepsis and Septic Shock.
  - But it ignores the context and is difficult to train given the sparsity
- The Paragraph Vectors embedding allows for a more specific training space, which is especially important in specialized domains such as medicine
  - The overall performance could reflect the bias and noise in the data, as more tuning would be able to further remove words that are not differentiating.
  - Upsampling sepsis samples would impact the embedding as they are not distributed across the vector space
- Using a CNN or Embedding layer improves training speed by quickly aggregating words and reducing the weight of non-relevant words. This is especially important with verbose text

Conclusion

- Both methods provide an improvement in accuracy in determining the outcome of a patient from standard clinical methods
- This ultimately shows the usefulness of free-from clinical text, which was previously excluded from automation systems

References

[1] Nicholas W. Sterling, Rachel E. Patzer, Mengyu Di, and Justin D. Schrager. Prediction of emergency department patient disposition based on natural language processing of triage notes. International Journal of Medical Informatics, 129:184–188, 2019.

[2] Alan Jones, Stephen Trzeciak, and Jeffrey Kline. The sequential organ failure assessment score for predicting outcome in patients with severe sepsis and evidence of hypoperfusion at the time of emergency department presentation. Critical care medicine, 37,5:1649–54, 2009.