Software Project: NLP to Save Lives? Improved Sepsis Prediction by Text Classification

Stefan Riezler

SoSe 2021

Sepsis

- Sepsis is a life-threatening organ dysfunction caused by an underlying infection, amounting to 20% of global deaths [Rudd et al., 2020].
- Sepsis is defined by thresholds on clinical measurements, e.g., the Sequential Organ Failure Assessment (SOFA) score [Vincent et al., 1996, Singer et al., 2016, Seymour et al., 2016].

Sepsis Prediction with Neural Nets

- Medical informatics has worked on machine learning tools for early prediction of severity scores, based on time-series of clinical measurements [Reyna et al., 2019].
- Inclusion of static information on pre-existing conditions has been shown to improve accuracy of machine learning prediction [Schamoni et al., 2019].

Improving Sepsis Prediction by NLP

- Anamnesis data is information about the medical history of a patient recorded in free-form text.
- Idea: Use text classification of anamesis data in a similar way as information on pre-existing conditions to improve sepsis prediction.

Project 1: Sepsis Prediction by Classification of Anamnesis Texts

- Stand-alone task
- Cleaning and normalizing of anamnesis texts (regex? seq2seq learning?)
- Dense representations of anamnesis texts (word-level? sentence-level? document-level?)
- Train classifier and use score as risk score (logistic regression?)

Project 2: Sepsis Prediction from Clinical Measurements

- Should not be stand-alone since not NLP
- Recreate prior work on neural networks for sepsis prediction (see [Moor et al., 2020] for an overview)
- Integrate Project 1 into neural network for sepsis prediction (Static feature? Ensemble? See also [Horng et al., 2017])

Data

Example data:

- https://www.cl.uni-heidelberg.de/statnlpgroup/teaching/softpro_sose2021/examples/
- Access to full dataset requires signing data protection statement.

References

Horng, S., Sontag, D. A., Halpern, Y., Jernite, Y., Shapiro, N. I., and Nathanson, L. A. (2017).
Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning.
PLOS ONE, 12(4):1–16.

Moor, M., Rieck, B., Horn, M., Jutzeler, C., and Borgwardt, K. (2020).
Early prediction of sepsis in the ICU using machine learning: A systematic review.

medRxiv.

Reyna, M. A., Josef, C. S., Jeter, R., Shashikumar, S. P., Westover, M. B., Nemati, S., Clifford, G. D., and Sharma, A. (2019).
Early prediction of sepsis from clinical data: The physionet/computing in cardiology challenge 2019.
Critical Care Medicine, 48(2):210–217.

Rudd, K. E., Johnson, S. C., Agesa, K. M., Shackelford, K. A., Tsoi, D., Kievlan, D. R., Colombara, D. V., Ikuta, K. S., Kissoon, N., Finfer, S., Fleischmann-Struzek, C., Machado, F. R., Reinhart, K. K., Rowan, K., Seymour, C. W., Watson, R. S., West, T. E., Marinho, F., Hay, S. I., Lozano, R., Lopez, A. D., Angus, D. C., Murray, C. J. L., and Naghavi, M. (2020).
Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the global burden of disease study.

The Lancet, 395(10219):200–211.

Schamoni, S., Lindner, H. A., Schneider-Lindner, V., Thiel, M., and Riezler, S. (2019).
Leveraging implicit expert knowledge for non-circular machine learning in sepsis prediction.
Journal of Artificial Intelligence in Medicine, 100:1–9.

Seymour, C. W., Liu, V. X., Iwashyna, T. J., Brunkhorst, F. M., Rea, T. D., Scherag, A., Rubenfeld, G., Kahn, J. M., Shankar-Hari, M., Singer, M., Deutschman, C. S., Escobar, G. J., and Angus, D. C. (2016).
Assessment of clinical criteria for sepsis for the third international consensus definitions for sepsis and septic shock (Sepsis-3).
JAMA, 315(8):762–774.

Singer, M., Deutschman, C. S., and Seymour, C. W. (2016).
The third international consensus definitions for sepsis and septic shock (Sepsis-3).
JAMA, 315(8):801–810.

Vincent, J., Moreno, R., Takala, J., Willatts, S., Mendonça, A. D., Bruining, H., Reinhart, C., Suter, P., and Thijs, L. (1996).
The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure.
Intensive Care Medicine, 22(7):707–710.