Master Thesis Proposal:

Early Prediction of Sepsis Onset through Time-Series Analysis of Vital Signs and Laboratory Parameters via Deep Learning Architectures

Sepsis is a life-threatening condition caused by the body's extreme response to an infection, often leading to organ failure and death if not treated promptly. Early and accurate detection is critical for improving patient outcomes. With advances in data science and deep learning, there is an increasing push to develop automated models capable of predicting sepsis. Recent studies have focused on leveraging electronic health records to build predictive models. These models typically employ techniques such as LSTMs and Transformers due to their ability to handle time-series data effectively.

The use of time-series data in predicting sepsis has gained increasing attention [1], [2]. Furthermore, studies integrating physiological time-series data and clinical notes [3] have explored ways to improve the predictive power of these models. However, the application of advanced deep learning techniques such as Transformers and LSTMs for sepsis prediction remains an ongoing area of exploration. Despite the availability of various prediction models, there is a significant gap in achieving high accuracy and reliability, particularly when relying on clinical measures. While advancements in machine learning and deep learning offer promising solutions, there is still a need to refine and optimize these models to handle complex and dynamic time-series data effectively.

The motivation for this research lies in the need to bridge this gap. By focusing on time-series vital signs and lab results and employing state-of-the-art deep learning models like LSTMs and Transformer-based architectures, this thesis aims to contribute to early detection and improved patient care. Current approaches have made significant strides, yet many existing models are not fully optimized. Many predictive models rely on isolated clinical features, which may not capture the complexity of sepsis progression. Moreover, there is a lack of exploration regarding the full potential of Transformer-based models in sepsis prediction, especially when combined with vital signs and lab results.

Research Questions

1. Can LSTM and Transformer-based deep learning models effectively predict the onset of sepsis using time-series vital signs and lab results?
2. What are the key features (e.g. vital signs, lab results) that contribute most to the accurate prediction of sepsis onset?
3. How do Transformer-based models compare to LSTM models in terms of prediction accuracy and interpretability for sepsis detection?

To address these questions, the thesis will involve the following steps:

1. **Literature Review**: A comprehensive review of the current state-of-the-art in sepsis prediction using deep learning, focusing on time-series data.
2. **Data Collection and Preprocessing**: Preprocessing available datasets such as MIMIC-IV for modeling, including data cleaning, normalization, and feature engineering.
3. **Model Development and Evaluation**: Developing and training both LSTM and Transformer models on the prepared data.

4. **Interpretability and Clinical Relevance**: Implementing techniques for model interpretability (e.g., SHAP values) to identify key features driving predictions.
5. **Model Comparison**: Comparing the performance of LSTM and Transformer models and selecting the most optimal model for sepsis prediction.
6. **Conclusion and Future Work**: Summarizing the findings and suggesting avenues for further research.

Timeline

- **Month 1**: Conduct a detailed literature review on sepsis prediction using deep learning models (LSTMs, Transformers) and review available datasets (e.g., MIMIC-IV) for suitable features.
- **Month 2**: Data preprocessing, including cleaning, normalization and feature engineering of time-series vital signs and lab results.
- **Month 3**: Develop and train the LSTM model, followed by model evaluation and adjustments based on performance metrics.
- **Month 4**: Develop and train the Transformer model, followed by model evaluation and fine-tuning.
- **Month 5**: Compare LSTM and Transformer models, focusing on accuracy, interpretability and clinical relevance. Implement interpretability techniques.
- **Month 6**: Finalize the thesis with a comprehensive analysis, results discussion and conclusions. Prepare the thesis for submission.

By the end of the project, I aim to have developed a reliable deep learning model capable of predicting sepsis onset using vital signs and lab results, offering significant contributions to early detection and improving patient outcomes.

Supervising Professor: Prof. Dr. Björn Eskofier

Supervisor: Dr. Alexandros Tanzanakis

Student: Fatemeh Salimiboroujeni

Medical Supervisor: PD Dr. med. habil. Florian Putz

Matriculation Number: 23148706

Start - End: 01.07.2025 - 31.12.2025

References:

[1] Zhang D, Yin C, Hunold KM, Jiang X, Caterino JM, Zhang P. An interpretable deep-learning model for early prediction of sepsis in the emergency department. Patterns (N Y). 2021 Jan 19;2(2):100196. doi: 10.1016/j.patter.2020.100196. PMID: 33659912; PMCID: PMC7892361.

[2] Tang, Y., Zhang, Y., & Li, J. (2024). A time series driven model for early sepsis prediction based on transformer module. BMC Medical Research Methodology, 24, 23.

[3] Yuqing Wang, Yun Zhao, Rachael Callcut, Linda Petzold. Integrating Physiological Time Series and Clinical Notes with Transformer for Early Prediction of Sepsis (2022) arXiv:2203.14469v1 [cs.LG]