Bachelor Project

English title: Representation learning of multimodal data using multiview machine learning

Danish title: Representation learning af multimodal data ved brug af Multiview machine learning

Name: Simon Stenbæk Jensen
Period: 12.02.2024-16.06.2024, ECTS 15
Supervisors: Tommy Sonne Alstrøm (tsal@dtu.dk), Thea Brüsch (theb@dtu.dk)

Description:
As data collection becomes increasingly automatic and more complex, methods of utilizing the data must evolve as well. Learning from multiple representations of the same data with multiview machine learning is a way of taking advantage of the abundance of data generated. This project focuses on training a classification model from ECG data and text descriptions of the same patients to increase classification accuracy. The project will be based on a current state-of-the-art model and aims at reproducing the model's results and apply it to new data and test the effectiveness of using other methods for pre-processing data before training and testing.

Timeline:

Running deliverables and plans

Throughout the project: Writing the thesis continually throughout the project duration for keeping up to date with current work and progress.

Week 1-4: 12/02/2024-10/03/2024

Literature study of the main paper in focus, "ETP", as well as relevant literature for reproducing the model components. This literature includes previous works that are referenced in ETP. Setting up a "to-do" list with said models and work from previous literature which needs to be implemented.

Preliminary work in preparation of implementation, such as preparing git repository, setting up environments and loading the primary dataset, PTB-XL.

End of week 4: 13/03/2024

Milestone: Hand-in project plan.

Week 5-7: 11/03/2024-31/03/2024

Implementation of models for preprocessing. Namely, ResNet18 for encoding ECG signals, and ClinicalBERT for encoding text. Success criteria is producing embeddings for both types of signals. If there is extra time, implementation of linear projectors to produce the multi-modal embedding space as well.

Week 8-10: 01/04/2024-21/04/2024

Implementation of linear projectors if not already done. Implementation of training and test loop and comparing to the original work.

Week 11-13: 22/04/2024-12/05/2024

Testing the models on new datasets.

Implementation of testing loop for performing zero-shot classification and accuracy testing.

<u>Week 14-18:</u>

Assuming the ETP model is fully implemented and evaluated.

Investigating the text reports semantics using data augmentations, to understand how strong an effect the medical language semantics have on the model effectiveness. Augmentations can be implemented like the ones in:

https://doi.org/10.1007/s10489-023-04453-3

Investigating the ECG signals in the PTB-XL dataset by training on only the ones that have been reviewed multiple times by medical professionals.

Investigating taking advantage of shared and unique information from each data modality using the approach from:

https://proceedings.neurips.cc/paper_files/paper/2023/hash/6818dcc65fdf3cbd4b05770fb957803e-Abstract-Conference.html

Specific learning objectives:

The student can

- Explain general concepts on representation learning and feature extraction.
- Understand and implement multiview machine learning methods using Python.
- Apply zero-shot learning methods.
- Combine and modify machine learning models for separate downstream tasks.