Digital Health Center
by Hasso-Plattner-Institut

Chair Digital Health & Machine Learning

Contact Benjamin Bergner or Prof. Christoph Lippert

Building a Text Embedding Model for the Imaging Research Warehouse at Mount Sinai *(topic is already reserved)*

Short Description: Text embedding models transform words or sentences into vector representations. Those vectors can be used in deep learning tasks such as text category classification or named-entity-recognition. In this thesis, you will access the Imaging Research Warehouse (IRW) at Mount Sinai, a database of thousands of images and radiology reports, to build such a language model. The model will subsequently be used to improve and evaluate a search engine, which is part of the tooling to access IRW data. You should be familiar with deep learning basics and committed to create a useful research product from start to end.

Lifelong Learning on Big Medical Imaging Data

Current AI algorithms for detection and diagnosis support in radiology are static, i.e. they are only trained once on a selection of data. However, medical images are collected continuously in the hospital, which may improve data hungry deep learning models. In this thesis, you will develop a lifelong learning pipeline, which is able to continuously improve itself on new data without the need to retrain on old data. If you succeed, there is potential to experimentally deploy your solution in a hospital setting.

Detection of Brain Abnormalities in Multimodal 3D MRIs

In this thesis, you will have access to ~15.000 de-identified brain studies, each consisting of multiple image series (MRI T1, T2 etc.). Accompanied to these are radiology reports, which are labeled normal/abnormal + types of abnormality. The target is to build a deep learning model that is able to classify these 3D scans and identify the location of abnormality. The challenge of this thesis is that the images are weakly labeled, i.e. there are neither bounding boxes nor segmentations, only classification labels are available.