Learning representations from time series data through contextualized LSTMs
WiML 2016 abstract

Time series are prevalent in biomedical data, from high-frequency vital signals, motion tracking waveforms and activity monitoring to longitudinal health indicators for patients with slow-progressing conditions such as osteoarthritis or cerebral palsy. The typical approaches to handle this data are either (1) too generic, not accounting for the characteristics (domain, mode of collection, context) of the dataset or (2) highly tailored to the dataset through painstaking feature engineering, requiring the intervention and supervision of domain experts in an iterative process. The former category includes tools such as PCA, FFT and embeddings via basis functions. An example of the latter is the processing performed on kinematic waveforms for gait analysis, which is impossible without a keen knowledge of biomechanics. While the former techniques are limited by their insufficient use of context, the latter are usually not generalizable, needing fine-tuning even for datasets of the same type.

We enable the use of context (structured covariates) to learning time series representations, thus obtaining a more generic framework for handling such data. This allows the incorporation of domain expertise in feature learning, flexibility in adapting to other datasets and easy integration of time series data with other data types such as structured covariates, text and images. By using multi-resolution LSTMs, we construct salient representations without the need for feature engineering.
Instead of deriving features for a given time series, the algorithm learns how to derive features for any number of temporal datasets, based on examples. We introduce two ways of attaining applicability for new types of data. First, we make the lower layers of the deep architecture flexible at runtime. This allows capturing the specifics of the data early in the transformation, while the rest of the network operates on a higher-order abstraction. For instance, we might train a deep network on step counts from accelerometers, but also use it on raw triaxial data. Additionally, this allows us to easily account for subject-specific traits. A second way to widen applicability is to directly model the similarities between samples. We consider the context of the data, which can include information about the domain, mode of collection or structured covariates linked to the time series. Contexts from different samples are introduced into a CNN that encapsulates context similarity, trained by backpropagating the difference in labels. The time series from both training and target samples are transformed through a multi-resolution network using LSTMs and convolutional layers. The feature examples given for the training samples are also incorporated.

We apply the procedure to accelerometer data from the Osteoarthritis (OA) Initiative, obtained from 2000 subjects, for a monitoring period of 7 days. This data is provided in terms of activity counts. The goal is to identify which subjects are at risk for fast OA progression. Previous studies on this data have applied simple transformations to the time series such as histograms or spline basis representations. Our model, which includes stacked LSTM layers for the time series combined with clinical covariates, improves the classification performance from 60% to 73%.