Reducing Annotation Effort through Projection Retrieval in an Active Learning Setting

Many applications of decision support systems require that their users comprehend the generated labeling suggestions. This is not always easy to achieve, especially when the underlying data is highly multivariate and when the models used to process it are complex. We have recently developed methods that reveal informative low-dimensional axis-aligned projections that contain predictive structure in data, provided that such low-dimensional structures exist. We introduce an approach that recovers informative projections in the more challenging active learning setting. Our framework selects samples to be labeled based on the relevant dimensions of the current classification model, trained on previously annotated data.

The effort is thus shifted to labeling samples that specifically target performance improvement for low-dimensional models. For this purpose, we enhance standard active selection criteria using the information encapsulated by the trained model. The advantage of our approach is that the labeling effort is expended mainly on samples that benefit models from the hypothesis class we are considering. Thus, high accuracy is achieved faster than with standard sampling techniques, reducing the data annotation effort exerted by domain experts. An added benefit is that the compact models are available to experts during labeling, in addition to the full-featured data. The informative projections highlight structure that experts should be aware of during the labeling process. Moreover, our active learning framework selects the most controversial, most informative and/or most uncertain data yet unlabeled (depending on the selected sampling technique), presenting it to the human experts in an intuitive and comprehensible manner -- typically using 2 or 3-dimensional projections -- which further simplifies the annotation process.

One application that justifies the need for our approach, is patient monitoring in the Intensive Care Unit (ICU). The ICU patients are connected to monitors, which continuously track the variability of multiple vital signs over time, in order to issue alerts whenever any of the vitals exceeds pre-set control limits. Typically, such deviations indicate serious decline in patient health status. Many of the issued alerts, however, are artifacts, triggered by accidental probe dislocation or inaccuracies of the monitoring equipment. In order to reduce alarm fatigue in clinical staff, the ideal monitoring system would dismiss artifactual alerts on the fly and allow interpretable validation of true alerts by human experts when they are issued. The strenuous effort required for the preparation of comprehensive training set is often compounded by the sheer complexity of the involved feature space. Precious expert time would thus be spent primarily navigating the dimensions of the data to establish grounds for labeling specific instances. Our system resolves this issue by determining a minimal set of unlabeled data for human adjudication, concurrently presenting the informative projections of this otherwise high-dimensional data.

Experiments show that this results in an improved learning rate over standard selection criteria, both for synthetic data and real-world data from the clinical domain. Additionally, the domain experts benefit from the availability of informative axis-aligned projections at the time of labeling. This comprehensible view of the data supports the labeling process and helps preempt labeling errors.