# Extbf{background integration of structure (both data- and expert-driven)

extbf{Background and Objective: }One of the most important hospital quality indicators is hospital readmission rate. The main reason is that unplanned readmissions have high human and financial consequences.

This motivated a number of models that are trying to predict readmission as a single label, binary problem (patient will be readmitted or not). However, such models do not answer the question extit{“ why is a patient at risk for readmission”}. Our objective is to develop interpretable predictive models that can explain why patient is likely to be readmitted (with which diagnoses) which will help making shift from predictive to prescriptive exploitation of predictive models in real medical practice. extbf{Method:}In order to address this problem we utilize machine learning method called Predictive Clustering Trees which is used for multi–label and hierarchical multi-label classification. Both (multi-label and hierarchical multi-label) area decision tree-like models who enable, not only prediction of readmission risk, but also prediction of the causes for potential readmission. In this case, we predict diagnoses and symptoms that are likely to appear on readmitting for specific patient.

For the hierarchical multi-label classification task we inspect the influence of data-driven and expert-driven hierarchies on predictive performance. Experiments are conducted on pediatric population of patients from state of California. extbf{Results:}The results of predictive models are thoroughly evaluated using example-based, label-based and ranking-based performance measures. Additionally, results are interpreted and discussed by medical doctors. Usage of structure improved accuracy from 21. 9\\% to 65. 2\\%.

Drastic improvement in performance is visible in other measures such as precision, recall, ranking loss etc.

extbf{Conclusions:}It is shown that integration of structure (both data- and expert-driven) increases predictive performance compared to flat models and that multi-label classification models based on Predictive Clustering Trees allow building of accurate and highly interpretable solutions.