

This thesis presents an applied research for the needs of a company Datlowe, s.r.o. aimed at automatic processing of legal documents. The goal of the work is to design, implement and evaluate a classification module that is able to assign categories to the paragraphs of the documents. Several classification algorithms are used, evaluated and compared to each other to be consequently combined to obtain the best models. The outcome is a prediction module which was successfully integrated into the entire document processing system. Other contributions, along with the classification module, are the measurement of the inter-annotator agreement and introducing new set of features for classification.