The aim of the thesis is to evaluate Czech text classification tasks in the low-resource settings. We introduce three datasets, two of which were publicly available and one was created partly by us. This dataset is based on contracts provided by the web platform Hlídač Státu. It has most of the data annotated automatically and only a small part manually. Its distinctive feature is that it contains long contracts in the Czech language. We achieve outstanding results with the proposed model on publicly available datasets, which confirms the sufficient performance of our model. In addition, we performed experimental measurements of noisy data and of various amounts of data needed to train the model on these publicly available datasets. On the contracts dataset, we focused on selecting the right part of each contract and we studied with which part we can get the best result. We have found that for a dataset that contains some systematic errors due to automatic annotation, it is more advantageous to use a shorter but more relevant part of the contract for classification than to take a longer text from the contract and rely on BERT to learn correctly.