

This thesis deals with anomaly detection of log data. Big software systems produce a great amount of log data which are not further processed. There are usually so many logs that it becomes impossible to check every log entry manually. In this thesis we introduce models that minimize primarily count of false positive predictions with expected complexity of data annotation taken into account. The compared models are based on PCA algorithm, N-gram model and recurrent neural networks with LSTM cell. In the thesis we present results of the models on widely used datasets and also on a real dataset provided by HAVIT, s.r.o.