This thesis focuses on automatic searching for clusters of topically similar texts in large text collection. We introduce an algorithm for nding the clusters and a method of optimizing its parameters using machine learning techniques. The algorithm is implemented and experimentaly evaluated. For evaluation we use a manually annotated collection of Czech documents, which contains a set of sample clusters chosen and tagged by a human annotator, and a huge collection of newspaper arcticles. Experiments show that the output of our algorithm ful ls our expectation and gives clusters of topically similar texts.