Automated Ham-Spam Lexicon Generation Based on Semantic Relations Extraction

AMIRA MOHAMMED IBRAHIM IDREES ,Ayman E. Khedr, Essam Shaaban

Abstract

One of the current essential methods for communication is electronic email (e-mail). It is currently considered the official method for different business activities such as conducting agreements, the setup of official meetings, and team collaboration. This continuous interest in e-mails as a communication channel has drawn the attention to the need for eliminating spam which have a vital effect on both network resources and business activities. This research focuses on generating a ham-spam lexicon based on text analysis which is aimed to be one of the main resources for detecting personal spam e-mails. The lexicon generation is a key step to efficiently and economically successful spam elimination. The proposed framework has proven its applicability on a dataset of six groups and the classification algorithms have been examined to prove the efficient classification. The research is a step in a wider view for general intelligent business communication and collaboration framework.