

論文の内容の要旨

論文題目 Minimally Supervised Approaches to Emotion Classification
using Unlabeled Data

(ラベルなしデータを用いた情緒分類への弱教師ありアプローチ)

氏 名 任 勇

In sentiment classification, conventional supervised approaches heavily rely on a large amount of linguistic resources, which are costly to obtain for under-resourced languages. Firstly we exploit the use of label propagation (LP), which is graph-based semi-supervised learning (SSL), to overcome this scarce resource problem.

Especially, we study several fundamental and critical issues such as controlling label propagation, choosing the initial seeds, selecting edges in the application of graph based SSL. Our evaluation on three real datasets demonstrates that manipulating the label propagating behavior and choosing labeled seeds appropriately play a critical role in adopting graph-based SSL approaches for this task.

Then we explore the identification of horror episodes, which is helpful to avoid triggering bad experience to users. We formulate horror recognition as a binary classification problem. Especially, we address it in the minimally supervised scenario where only a handful of labeled samples are available. Besides adopting conventional text features including unigrams and bigrams, we attempt to exploit features derived from a large scale unlabeled corpus in this task. We investigate the usefulness of the induced features through a series of experiments, using newly emerging literature, hint fiction, as typical, concise example of horror episodes.